





王楗楠 王洪波 编著

- 从繁杂的SEO知识体系中，**提炼SEO操作方法和SEO思维**
- 通过实例教学**进一步完善对SEO的理解和认识**
- 通过经典案例的分析，**掌握SEO的应用技巧与方法**
- **即刻让你的网站流量与人气倍增长**

清华大学出版社

SEO 网站营销推广全程实例



王楗楠　王洪波　编著

清华大学出版社
北　京

内 容 简 介

本书详细介绍了 SEO 优化的实际应用技术，深度剖析了搜索引擎的工作原理、关键词研究、代码优化、网站构架优化、链接建设及策略更正。此外，书中还介绍了作弊与惩罚、排名等因素，以及常用 SEO 工具的使用。本书不仅内容丰富完整，更重要的是本书是笔者以多年 SEO 实践经验根据网站优化实施的流程及步骤制定的。有些人认为 SEO 只意味着发垃圾和欺骗性的手段，这是不正确的。SEO 可以为用户设计一个清晰易用的网站，也可以让计算机通过跟踪链接发现新的页面。关注 SEO 可以发现用户寻找产品或网页时输入的关键词，然后就可以在页面上自然融入这些词。

本书不仅对 SEO 初学人员有帮助，对进行网站运营、建设、推广、设计的网站站长、网络技术营销团队及传统商务公司经理等都有参考价值。

图书在版编目（CIP）数据

SEO 网站营销推广全程实例/王楗楠,王洪波编著. —北京：清华大学出版社，2013.1
ISBN 978-7-302-30625-2

Ⅰ.①S…　Ⅱ.①王…　②王…　Ⅲ.①网络营销　Ⅳ.①F713.36

中国版本图书馆 CIP 数据核字（2012）第 271727 号

责任编辑：夏非彼
封面设计：王　翔
责任校对：闫秀华
责任印制：王静怡

出版发行：清华大学出版社
网　　址：http://www.tup.com.cn，http://www.wqbook.com
地　　址：北京清华大学学研大厦 A 座　　**邮　　编：**100084
社 总 机：010-62770175　　**邮　　购：**010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质量反馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
印 装 者：清华大学印刷厂
经　　销：全国新华书店
开　　本：190mm×260mm　　**印　张：**15　　**字　数：**384 千字
版　　次：2013 年 1 月第 1 版　　**印　次：**2013 年 1 月第 1 次印刷
印　　数：1~4000
定　　价：32.00 元

产品编号：047481-01

前　言

SEO，全名为 Search Engine Optimization，翻译为搜索引擎优化。其主要工作是通过了解各类搜索引擎的各项技术对网页进行相关的优化，使其提高搜索引擎排名，从而提高网站访问量，最终提升网站的销售能力或宣传能力的技术。

本书的特色：

（1）从纷繁复杂的 SEO 知识体系里，提炼了 SEO 操作和 SEO 思维两个部分。

（2）通过实例教学，使读者对 SEO 的理解和认识更加完善，从而进一步对 SEO 知识体系去繁存真。

（3）本书关键在于培养读者的思考方式，这才是能否成功的关键。

基于以上这些思路，笔者探索编写了这本 SEO 网站营销推广全程实例。本书各章节的内容如下：

第 1 章介绍了 SEO 的基础知识。SEO 搜索引擎优化的主要任务就是通过使用 SEO 技术让网站的关键字在百度或谷歌等搜索引擎中排到好的位置，使之能够自然性地为我们的网站带来大量的廉价流量；通过 SEO 为自己的网站带来了人气，使其能够让自己的网站营利。上面都是一些通俗的解释，专业的描述，SEO 是指通过架设符合搜索引擎索引排序算法的网站结构、内容、代码以及外链达到预先设定的排名效果，并通过排名获取流量来达到网络营销的目的。

第 2 章介绍了搜索引擎的基础知识。说到搜索引擎大家都知道搜索引擎是个效劳器，并且也了解搜索引擎是分为两个主体，一个是用来统计数据的搜索引擎效劳器；另一个则是收录页面信息的搜索引擎机器人，也就是蜘蛛，这两种缺一不可。本章节主要为读者介绍搜索引擎的原理及工作流程。此外还将讲述搜索引擎各主要工作与搜索引擎优化之间的对应关系，为学习搜索引擎优化提供了理论依据，并且举出几种搜索引擎可能用到的排名计算方法。

第 3 章主要讲解了关键词的概念。想要对网站进行搜索引擎优化，肯定先要为网站的页面选择适当的关键词，之后围绕着核心关键词展开相关的工作以使关键词能在搜索引擎排名中到达满意的结果。而很多人在选择关键词的时候，往往选择一些热度非常高，搜索量非常大的词来做关键词，吸引流量。这个方法无可厚非，但是，使用这种方法也是要讲究方法和原则的。我们选择热度关键词的时候，不要脱离了自身网站的主题内容。

第 4 章描述了 URL 网页地址优化的办法。如今随着互联网的盛行，网站建设、网站优化也出现了很多的问题，比如建网站不规范或者后期的网站不规范，最为严重的可能就导致网页 URL 标准化问题。本章节就重点介绍如何优化网页 URL 的问题，介绍的内容包括 URL 网页地址优化的基础知识，网页地址的命名方法，分隔符的使用，网页的地址长度定义，以及

网页地址重定向和静态化的问题。

第 5 章介绍页面代码优化知识。页面是网站建设中的基本组成部分，代码优化则是网站页面优化中的一个重要环节，也是页面优化的基础，SEO 搜索引擎优化要做的基本工作就是改善网页代码。本章将具体介绍页面的代码优化，主要包括页面代码的精简方法，HTML 标签优化，标头优化，权重标签的使用，图片的优化以及沙盒效应知识。

第 6 章讲解网页组成结构。要做好网站的内部优化就需要了解网站的组成结构等，需要对网站整体构架有个全面的认识。本章主要介绍网页的组成结构及对优化的影响，包括网页结构简介的基础知识，网页组成的基本概念。

第 7 章描述了网站结构优化。目前已经很少有网站是由单独的页面组成，一般网站都会有几个目录（频道）、无数网页。对搜索引擎优化而言，当涉及多个尤其是成千上万个网站页面时，往往就需要有清晰的网站结构，以确保搜索引擎和用户可以方便地进行访问。而网站结构的规划，对各级关键词优化力度的分配、网页权重的传递都起着非常重要的作用。

第 8 章讲解链接策略知识。链接（即超链接）指从一个网页指向另外一个目标的连接关系，这个目标可以是一个网页，也可以是同一网页的不同位置，还可以是一个图片、电子邮件地址或文件，甚至是一个应用程序。在网站优化工作中，链接的优化也很重要。本章就介绍链接的策略知识，包括站内链接的优化方法，内部链接的优化，外部链接的优化，增加外部链接的办法。

第 9 章讲述关于优化误区与作弊的基础知识。搜索引擎优化的作弊是针对于搜索引擎算法的不全面而采取对应的欺骗性手段，提升了页面权重及相关性的行为，按照作弊的范围，可以划分为内部作弊以及外部作弊。本章就重点介绍内部作弊和外部作弊的一些基础知识，防止在实际的优化工作中重蹈覆辙。

第 10 章介绍一些 SEO 辅助工具。SEO 辅助工具就是对搜索引擎功能优化处理的工具。用搜索引擎检验网站的搜索引擎优化状况是最直接、最有效的方式，因为搜索检索结果可以带给人们大量有价值的信息。通过搜索引擎检索结果的分析，是研究网站搜索引擎优化状况的有效方法之一。因此搜索引擎优化工具，又称为 SEO 优化工具，其通过检查网站链接数量、网站被搜索引擎收录网页数量、网站的 PR 值等。它只能在一定范围内反映出某些指标的状况。应用搜索引擎优化工具的目的，是为了检验一个网站对搜索引擎的友好程度及其可能获得的搜索排名效果。

第 11 章介绍了经典案例分析的办法。分析别人的网站也是学习 SEO 的技巧之一，通过分析竞争对手的网站来选定自己网站的优化方向是很多站长采用的方法，也是相当行之有效的。对排名较好的网站进行细致地 SEO 分析绝对会给你带来意外的好处。本章选一些比较有代表性的网站来分析其中优化的利与弊。

本书由 SEO 网站优化实战经验丰富的资深设计师编写，理论、实践相结合，章节安排合理，注重读者的实用性、可操作性。全书由王楗楠、王洪波主编，另外韩宇轩、吴海娇、崔健、杨雪、秦树德、张铭运、赵红、陈章、于海鑫、任霖等参与了本书部分内容的编写工作。由于作者水平有限，加之创作时间仓促，本书疏漏之处在所难免，欢迎各位读者与专家批评指正。

编者

2012 年 7 月

目 录

第 1 章 SEO 简介

第 2 章 搜索引擎

第 3 章 关键词

第 4 章　URL 网页地址优化

第 5 章　页面代码优化

第 6 章　网页组成结构

第 7 章　网站结构优化

第 8 章 链接策略

第 10 章　SEO 辅助工具

第 11 章　经典案例分析

第 1 章　SEO 简介

SEO 搜索引擎优化的主要任务就是通过使用 SEO 技术让网站的关键字在百度或谷歌等搜索引擎中排到较好的位置，使之能够自然地为我们的网站带来大量的廉价流量。这里流量的概念是：如果我们把网站比作一个超市，那么流量就相当于来逛超市的顾客。这种在网络上的顾客，就被称之为 IP 量。通过 SEO 为自己的网站带来了人气，使其能够让自己的网站营利。以上都是一些通俗的解释，专业的描述法，SEO 是指通过架设符合搜索引擎索引排序算法的网站结构、内容、代码以及外链达到预先设定的排名效果，并通过排名获取流量来达到网络营销的目的。

1.1　互联网营销

先来说说互联网营销，也可叫做网络营销，它是以网络为基础，使用数字化的信息和网络媒体的交互性来辅助营销目标得以实现的一种新型的市场营销方式。

互联网营销有着很强的实践性特点，从实践中发现互联网营销的运作方法和规律，比浮泛的理论讨论更有实际意义。“中国网络营销网”汇集了最新、最全、最完善的网络营销资讯。因此，如何定义网络营销其实并不是最重要的，关键是要理解网络营销的真正意义和目的，也就是充分熟悉互联网这种新的营销环境，利用各种互联网工具为企业营销流动提供有效地支持。这也是为什么在网络营销研究中必须要正视网络营销实用方法的原因。

随着网络技术的发展和成熟以及网络运作成本的低廉，如今的网络就好比是一种“粘合剂”，将各种企业、团体、组织以及个人跨时空的连接在一起，使得他们之间信息的交换变得唾手可得。市场营销中最重要也最本质的是组织和个人之间进行信息传播和交换的过程。如果没有信息交换，那么交易也就是无本之源。正因如此，互联网具有营销所要求的某些特性，方使得网络营销呈现出多种特点。目前对于网络营销的特点，学术上有两种比较主流的观点，事实上都大同小异。

1.2 搜索引擎营销 SEM

SEM 搜索引擎营销，包括 SEO 和 PPC 两部分。主要是利用搜索引擎的关键词排名带来的流量进行宣传推广。其中 SEO 的全名为 Search Engine Optimization，翻译为搜索引擎优化。但有时 SEO 也指 Search Engine Optimizer，即搜索引擎优化师。而 PPC 的全名为 Pay Per Click 可译为点击付费竞价广告，其典型代表有 Google Adwords 广告，以及百度竞价排名。

优化后的网站将不仅仅在 Google 和百度中提高排名，而且排名的提高将影响到世界任何一个搜索引擎。

1.3 关键字竞价排名

关键字竞价排名的基本特点是需要按照被点击数量付给搜索引擎一定的费用。推广信息泛指在搜索结果中，一般是靠前的位置在右侧会出现“推广链接”链接文字。如果用户没有点击此链接，则不会收取推广费用，如图 1-1 所示。



图 1-1　百度竞价排名

一般在搜索引擎营销使用中，竞价排名所起到的主要作用及其特点基本有以下几方面：

- 按效果付费，用度相对较低。
- 泛指在搜索结果页面，与用户检索内容高度相关，增加了推广的定位程度。
- 竞价结果泛指在搜索结果靠前的位置容易引起用户的关注和点击，因而效果比较明显。
- 搜索引擎自然搜索结果排名的推广效果是有限的，尤其对于自然排名效果不好的网站，采用竞价排名可以很好地弥补这种劣势。
- 企业可以自己控制点击价格和推广用度。
- 企业可以对用户点击情况进行统计分析。

竞价排名，是一种按效果付费的网络推广方式。用少量的投入就可以给企业带来大量潜伏客户，有效晋升企业销售额和品牌知名度。

使用竞价排名推广网站所搜引擎会按照给企业带来的潜伏客户访问数目来计费，网站管

理者也可以灵活控制网络推广投入的资金，意图获得最大回报。

竞价排名有以下几个优点。

- 见效快：充值后设置关键词价格后即可进入。
- 枢纽词数目无穷制：可以在后台设置无数的关键词进行推广，数目自己控制，没有任何限制。
- 关键词不分难易程度：不论多么热点的关键词，只要你想做，都可以进入前三名甚至第一名。

在对竞价排名有所了解之后，下面将进一步介绍搜索引擎优化的相关内容。

1.4 什么是 SEO

Search Engine Optimization 中文译为搜索引擎优化，简称 SEO，是目前最为流行的网络营销方式之一。主要目的是增加特定关键字的曝光率以增加网站的能见度，进而增加销售的机会。SEO 分为站外 SEO 和站内 SEO 两种。SEO 的主要工作是通过了解各类搜索引擎如何抓取互联网页面、如何进行索引以及如何确定其对某一特定关键词的搜索结果排名等技术，来对网页进行相关的优化，使其提高搜索引擎排名，从而提高网站访问量，最终提升网站的销售能力或宣传能力的技术。

1.4.1 站外 SEO

站外 SEO，也可以说是脱离自身网站站点的搜索引擎优化技术，命名来自外部站点对自身网站在搜索引擎自然排名的影响，这些外部的因素是超出自身网站的控制的。最有用、最强大的外部站点因素就是反向链接，即我们所说的外部链接。

毫无疑问，外部链接对于一个站点收录进搜索引擎结果页面起到了重要作用。下面来看如何才能产生高质量的反向链接。

（1）高质量的内容

获取高质量外部链接最有效的办法就是编写出高质量的文章，你的文章内容能够让读者产生阅读的欲望从而对文章进行转载。

（2）合作伙伴、链接交换

与合作伙伴互相推荐链接，与行业网站、相关性网站进行交换链接。

（3）分类目录

将网站提交到DMOZ目录、Yahoo！目录、ODP 目录等专业目录网站。

（4）社会化书签

将网站加入百度收藏、雅虎收藏、Google 书签、QQ 书签等社会化书签。

（5）发布博客创建链接

目前获取外部链接最有效的方式之一就是通过博客发布文章。

（6）论坛发帖或签名档

在论坛中发布含有链接的原创帖或者在编写的签名档中插入网址。

（7）购买高价值链接

个人不建议使用此方法，如果被搜索引擎发现则会被降权。

（8）与 SEO 业务合作伙伴进行 SEO 内部优化

1.4.2 站内 SEO

站内 SEO 就是通过对网站内部的调整而起到对搜索引擎友好的目的。网站想提高关键字排名，做好内部工作是必须的，因为搜索引擎优化是个系统的工程，不是一蹴而就的，需要大量的积累资源和尝试修改。网址内部优化主要有以下几方面。

（1）网站内部的链接结构

尽量改变原来的图像链接和 Flash 链接，使用纯文本链接，并定义全局统一链接位置。

（2）标题 title 的重新定位

标题中需要包含有优化关键字的内容，同时网站中的多个页面标题不能雷同，至少能显示“关键字、网站首页及一段简单的含关键字的描述”类型。标题一旦确定就不要再做修改。

（3）每个页面包含有关键字并保持一定的频率

对内容结构做了简单的调整之后，立即到搜索引擎登录，希望能尽早收录新标题和新描述。

（4）网站结构做细节调整

假设因为原有的网站是形象页面，运用了比较多的 Flash 和图像，这些页面元素不利于搜引擎的收录，所以需要在该网页的下方加入 3 栏，分别是相关的公司简介、关键字产品、新闻与公司的关键词产品列表，并对这 3 栏内容添加 URL。

当然，最好的方法是使用新闻或资讯栏目更新关键词产品新闻。可以制作一个从首页链接跳转至一个单页面作为关键字的详细描述。该页面的描述内容包含了公司关键词产品列表链接。这些都是为了形成企业站点内的网状结构。

（5）资源应用

将网站结构大致调整好了之后，就可以利用资源进行外部链接的扩展了。首先是可以使用百度空间，空间的域名就使用公司产品的关键字，同时进行公司原网站信息的转载，附带

公司网址，让百度 robot 在第一时间访问本站点。其中还有一个技巧：使用该空间账号去随机访问百度空间内的其他用户，以获得回访，这样蜘蛛到达的效果会更好。



在百度空间、贴吧和知道发表的时候记得要附带链接信息，方便互访，以提高访问量。

SEO 是一项全面细致的工作，如图 1-2 所示。

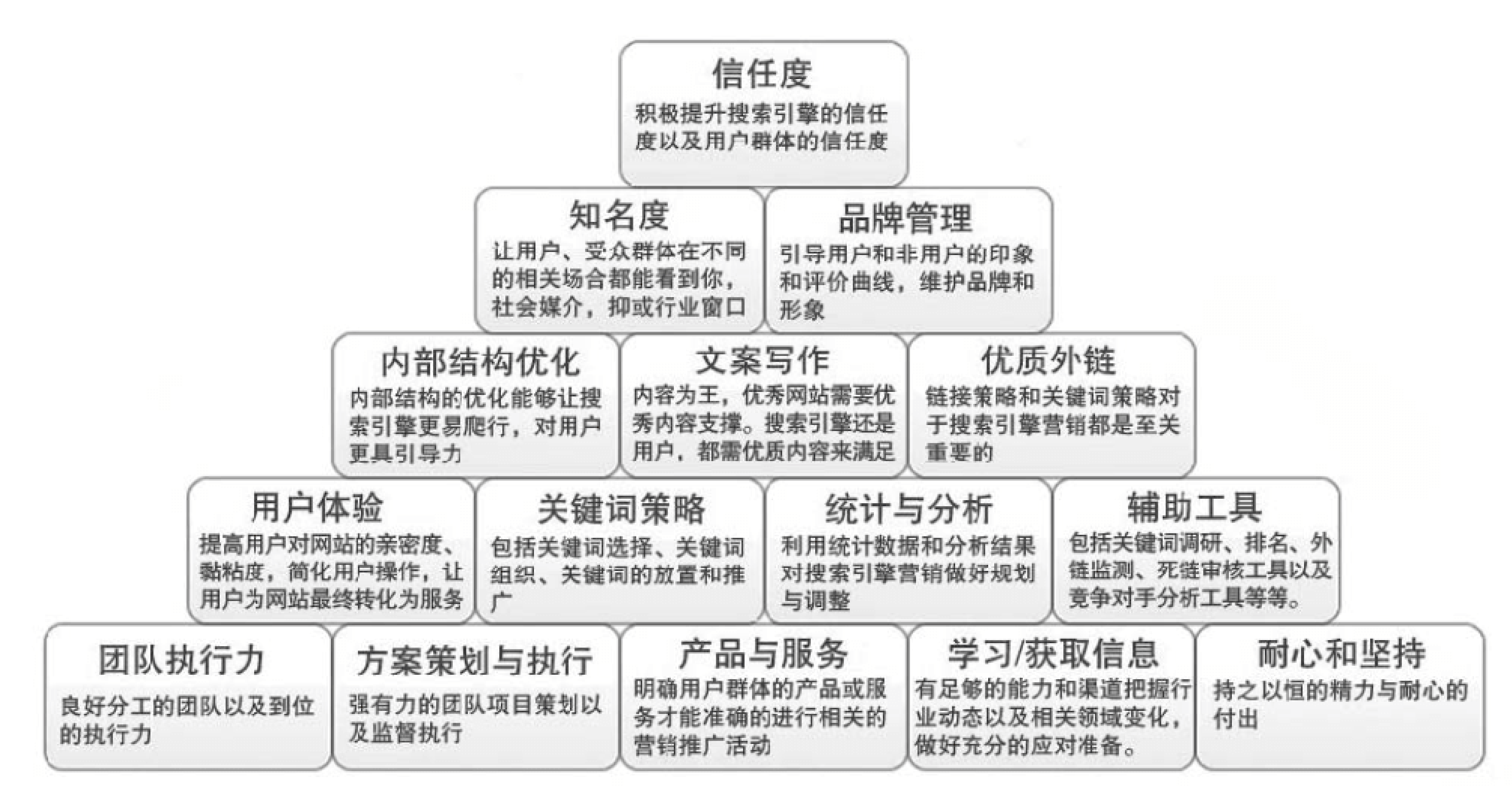


图 1-2　SEO 工作全解图

1.4.3　SEO 与竞价排名有何不同

所有优化人员都知道的两种最基本的网站推广方法就是竞价推广和 SEO。竞价就是搜索引擎竞价排名，即是竞价排名网站（不仅仅包括搜索引擎，还包括一些电子商务站点）推出的网站推广增值业务。流量越多，知名度越高。

当前几乎所有搜索引擎都有推出自己的推广产品服务，竞价排名推广包括我们所熟悉的百度竞价推广和 Google Adword。它们主要是以出售相关关键字给予用户相应的竞价排名的方式获取利润，用户也通过购买与自身产品或服务的相关关键字进行推广销售来获取利益。在百度、Google 里搜索某个词语，出来的十几个结果页中，右下角显示“推广”字样的就是通过给百度、Google 一定的广告费用的“竞价推广”，如图 1-3 所示。

用户在搜索引擎里输入某个词语，可以反映出他的某种需求和心态，这里的竞价推广就是巧妙地利用了这一点，所以，当我们搜索某个词语时，出现了很多排名靠前的推广链接就是如此。它有效的将有某种需求的用户与产品服务提供方结合在一起，其作用就是可以有效地推广、销售产品或服务。

图 1-3　百度搜索关键词“婚纱”所出现的竞价推广链接

SEO 是一门搜索引擎优化技术，简单地说，就是网站优化技术。SEO 优化出现得比竞价排名要早。一个推出竞价排名的网站肯定是“痛恨”SEO 优化的，但又是感恩 SEO 优化的，否则他们还想不到用竞价排名来赚钱的招。也许在很久很久以前（百度还在专心研究搜索引擎抓取技术而非营利手段的时候），就有这么一个人，他很喜欢在百度上东搜搜，西搜搜。后来他在想，这搜出来的东西，要都是自己想要的结果那该多好。其实不光是他，所有用搜索引擎的人都有过这样的想法。SEO 就是通过优化某个有价值的关键字获取搜索引擎排名进而推广、销售某产品的服务。当然不管是竞价推广还是 SEO 优化，其实都是网上赚钱的一种手段。只要你有好的产品、有价值的关键字，就可以利用竞价、SEO 赚到钱，不过笔者更青睐利用 SEO 赚钱，因为它是免费的而且长期有效。

1.4.4　SEO 的发展变化

到了现在，搜索引擎的算法更加的复杂，考虑的问题也增加了很多，已经不仅仅是外部链接或者关键词因素了。随着微博以及 SNS 的发展，在以后的算法中还有可能会涉及到众多其他的因素。现在的书签、网摘、用户点击率及网站跳出率等都是排名的一个因素，所以，如果想要网站有良好地发展，首先要做的事情就是有丰富的内容以及良好的用户体验。

大体上来说，搜索引擎优化的目的与搜索引擎的目的是一样的。优化是为了能让我们的网站有更好的用户体验以及获得更多的流量，而搜索引擎是为了提供用户找到他们想要的信息。对于做 SEO 的人来说，最重要的是提供一些有价值、与网站主题相关，最好是高质量的原创文章。而搜索引擎要做的就是通过自己的算法，把所有符合条件的内容，进行多种计算，最终返回价值高、对用户有帮助的内容展现给用户，所以 SEO 和搜索引擎是相互并存的。

目前搜索引擎优化在国内流行的时间并不是很长。而就在 SEO 刚刚兴起的前期，站长们最通常使用的手段就是关键词堆积，并且这个方法在当时非常有效。在标题标签、关键词标签或者在网页正文中放上大量的目标关键词，过几天就会有不错的排名和流量。很多站长甚

至会在网页中放上很多与内容不相关的词，从而实现从搜索引擎截获大量流量。但是好景不长，当很多站长都开始用这种方法疯狂地堆积关键词或者滥用标签，这个时候搜索的质量肯定就会下降，也会直接影响到搜索引擎。当初最知名的搜索引擎 Alta Vista 就是因为这样被用户弃用的。所以，当很多站长在钻空子时，搜索引擎也会进行相应的调整，对这些作弊的手法惩罚得比较严厉。

目前随着关键词的堆积不再有效果，而开始进入了全新的 SEO 时期，那就是外部链接的建设。谷歌发明了网页级别 PR 以及百度的超链接算法等都是以外部链接为基础判断的，简单来说，外部链接越多，就意味着 PR 值越高，权重越高，所以网站的关键词排名也就会越高。很多站长都走上了自己疯狂的外部链接之路。不论是论坛发帖，还是大量制造垃圾链接，购买链接，或者制造站群，这些都对排名起到了作用。但是后来搜索引擎也开始进行反击，不论是 Google 还是百度，都制定了相应的对策比如 Google 的沙盒，百度的新站观察期等，这些对策都能让外部链接起到减缓的作用。因此很多站长坚持了一段时间后看到没有希望就放弃了。

使更多人了解搜索引擎的运行原理，引导大家广泛合理的 SEO 行为，让认真做原创优质内容的创业者能够得到更多的流量，令抄袭、抓取别人内容的建站行为得到警惕，是百度作为中文搜索领域的领导者应有的责任和义务。只有这样才能有效支持互联的网创新力量，使互联网生态圈得到更加健康有序的发展。

1.5 SEO 的利与弊

SEO 优化的自然排名是搜索引擎固有的排名，是基于搜索引擎算法的自然排列结果。在各种网络营销手段中，SEO 是性价比最高的营销手法：成本低、针对性强，用户认可度高。搜索引擎优化 SEO 是电子商务平台，企业网站等各种以营利为目的，或是宣传推广信息、提高品牌知名度，宣传公司品牌的网站都必须使用的一种营销手法。下面来分析 SEO 的利与弊。

1. 自然排名优势

首先自然排名一般比竞价排名的点击率更高，其次自然排名给人一种品牌的效果，更容易给人信任感。排名稳定持久，从长远看，自然排名优化费用远远低于竞价排名费用，相比其他网上网下营销形式，费用低，投资回报率高。自然排名比竞价排名更令搜索用户信任，获得的点击量是竞价排名的两倍以上。

2. 自然排名的劣势

排名不是自动的，需要网站管理人员的搜索引擎优化以及内容的提升。自然排名是需要时间才能出效果的，而竞价排名的速度却是很快的，如谷歌 5 分钟就可以搞定竞价排名了。

3. 竞价排名的优势

充值后再设置关键词价格即可进入百度排名前十，位置可以自己控制。关键词数量无限制：可以在后台设置无数的关键词进行推广。关键词不分难易程度：不论多么热门的关键词，只要你想做，都可以进入前三甚至第一。

4. 竞价排名劣势

不继续支付费用排名会立即消失，关键词竞价的费用越来越高，长期看护费用不低，相互竞价导致每次点击价格窜升，存在欺骗性点击，使广告预算增加。竞价排名不仅取决竞价，还与点击量有关，如果竞价广告不能产生高点击率，可能会被终止。目标用户对广告有天然的屏蔽习性，更愿意点击自然排序。

如果网站想要长期稳定的发展，并且想让网站给企业带来长期的效益，那么网站应该在建站的初期就开始考虑优化的发展方向，这也是目前最常用的为网站带来流量的网站策划之一。

- 专业的搜索引擎优化可以提升企业网站的形象。
- 企业网站的关键字完全可以通过搜索引擎优化来做，而且同样可以做到非常稳定，效果明显。
- 关键词增值效益，专业的搜索引擎优化可以把初期做上去的几个关键词扩展到几十甚至几百个关键词，那么后期的效益就不言而喻了。

只有这样的网站才是一个企业所真正需要的，而不是短期的百度竞价排名带来的效益能够相提并论的。

1.6 SEO 的应用领域

围绕着搜索引擎排名，网站应站在利于引擎沟通友好的立场来优化网站。多年的网站运营经历让我们对 SEO 有了更深地了解。在跟客户交换时，还能了解到 SEO 不仅是引擎的优化，而是一个网站运营的优化。所以应当发扬这类优化，来改造我们的互联网，给那些依托互联网的人缔造一个杰出的空间。

SEO 基础知识、SEO 应用领域和 SEO 六种高质量外链等的网络推广能获得很好的 SEO 排名，引擎优化被各行各业的网站普遍利用，下里是常被利用的三个范畴。

1.6.1 企业网站

企业网站就是企业在互联网上进行网络建设和形像宣传的平台。企业网站就相当于一个企业的网络名片，不但对企业的形象是一个良好的宣传，同时可以辅助企业的销售，甚至可以通过网络直接帮助企业实现产品的销售。企业可以利用网站来进行宣传、产品资讯发布、招聘等。企业网站的作用就是为展现公司形象，加强客户服务，完善网络业务，还可以与潜在客户建立商业联系。随着网络的发展，也出现了提供网络资讯为营利手段的网络公司，在这些公司的网站上提供人们生活各个方面的资讯：如时事新闻、旅游、娱乐、经济等。

在互联网早期，网站只能保存单纯的文本。经过多年的发展，当万维网出现之后，图像、声音、动画、视频，甚至是 3D 技术也开始在互联网上流行起来，网站也慢慢地发展成我们现在看到的图文并茂的样子。通过动态网页技术，用户也可以与其他用户或者网站管理者进行交流，还有一些网站提供电子邮件服务。

1.6.2 电子商务型网站

电子商务型网站进行优化后可以通过搜索引擎向更多的潜在消费者倾销本身的商品，从而节约巨额的广告费用，增加产品的销量。

电子商务型网站就是基于互联网的交易行为。一般包括 B2B、B2C、C2C 三类，分别指企业对企业（阿里巴巴、环球资源网），企业对个人（tmall、京东、国美网上商城等），个人对个人（淘宝网、拍拍网）。

其中 B2B 主要营利来源为会员制度和广告。B2C 主要的营利来源分平台型 B2C（tmall，将很多企业销售给个人）和独家型国美网上商城（一家企业销售给个人）。平台型的主要营利来自会费、分成和广告；独家型主要营利来自于产品、服务的销售利润差。C2C 就是个人销售给个人，跟独家型 B2C 类似。

1.6.3 内容型网站

大多数的网站，其本质上都是属于内容型网站。而很多本来是纯功能型网站的（如在线 RSS 浏览器、网摘等），由于功能无法直接收费，也走向了内容型网站的道路。资讯内容类网站优化后可以大大提高网站的流量，从而进一步吞食强大的竞争对手，最终后来居上，成为行业的领先者。如泡泡网、123 宠物网的强势发展，如图 1-4 所示。



图 1-4　123 宠物网狗市频道

1.7 SEO 工作内容

SEO 需要从网站的内容与网站的外部两方面同时开展各种不同的操作方法，以对搜索引擎友好。

SEO 优化前要做以下几个准备工作：

- 域名的选取，要易于客户记住，往往是越短越好。
- 一个稳定的服务器空间，便于客户访问，蛛蛛抓取。
- 良好的程序，代码的规范性，对蛛蛛的友好性。

1.7.1 SEO 的工作阶段

具体工作的五个阶段（可以算四个阶段）如下：

（1）网站的收录，有收录才可能有排名，没有收录一定没有排名。

（2）关键字排名的不断提升，依靠对网站内优外化的不断工作。

（3）关键字良好的排名带来的网站流量。

（4）对网站流量进行转化，直接或间接的带来效益。

（5）如果说有第五个阶段，那就是不断重复以上四个阶段，以让网站能长远发展，获得最大的收益。

之后便可以利用博客养站法、论坛养站法及利用收藏夹，对网站进行收藏；可以利用微博做网站的外链；可以去友情连接平台，发布站点信息；可以去找一些同行业的友情连接来交换；可以做一些问答，带去网站的名气；可以与一些资源共享平台共享一些含有网站信息的文件；可以与媒体合作，写一些与公司情况或产品情况相关的信息，去百度百科、SoSo 百

科、维基百科等写一些与公司相关的信息或产品信息、领导人员情况等，以增加曝光率；可以找一些导航站进行提交，收录等；可以去一些信息发布平台，发布一些相关网站的信息等。

作为一名 SEO 工作者每天应第一时间查看网站的收录，反向连接。打开网站查看网站是否正常运行；查看网站流量与访问量、IP、跳出率等；查看网站排名、网站转化率、更新网站内容、综合观察网站情况、及时调整网站方案和更新网站内容等。

1.7.2 SEO 核心宗旨

目前网站优化人员似乎都非常在乎搜索引擎蜘蛛的体验。而实际上，用户体验也是十分重要的，搜索引擎也是一直在模仿人们的搜索习惯。用公式概括如下：

```
SEO/SEM =用户体验+蜘蛛体验
```

这个公式可以称为 SEO/SEM 定理，公式中各个要素说明如下。

搜索引擎营销（Search Engine Marketing，简称 SEM），是一种新的网络营销形式。SEM 所做的就是全面而有效地利用搜索引擎来进行网络营销和推广。SEM 追求最高的性价比，以最小的投入，获得最大的来自搜索引擎的访问量，并产生商业价值。SEM 也包括竞价排名、购买搜索引擎竞价排名广告等付费手段。

用户体验（User Experience，简称 UE），是一种纯主观的在用户使用一个产品（服务）的过程中建立起来的心理感受。在早期的软件设计过程中用户体验主要是来自用户和人机界面的交互过程，现在流行的设计过程注重以用户为中心。用户体验设计目标是什么？最重要的是要让产品有用，这个有用是指用户的需求，其次是易用。具体到网站优化 SEO 上，一个网站用户体验好，就是说用户在网站上很容易找到所需要的信息，并且没有影响用户浏览信息的障碍物，如弹窗、广告等无用信息。

网页载入时间都是一个反馈用户体验的指数。网页载入速度越快，对网页权重提升都有极大的帮助。

蜘蛛体验（SEO Experience，简称 SE），是搜索引擎算法对网站的评价，算法对网站评价好则排名就好。算法是怎样评价网站的呢？大量优质外链、原创性文章、高频率更新、网站结构算法好，用户体验等符合算法并对网站评价就好。

证明：

用户体验是如何提高 SEO 排名呢？其一，好的用户体验能提高网站 PV 值产品转化率。最重要的是把流量变成用户，用户日积月累的流量就越大。

假设有两个网站 A、B，它们除了流量不一样外，其他参数都相同。

A 的流量 > B 的流量，那么 A 的权重与排名就优于 B。

即：用户体验好的网站权重与排名好。

遗忘 SEO，记得某个资深人士也曾经说过相似的话。遗忘 SEO 就是指不要为了 SEO 而去做 SEO 优化。SEO 思想必须要贯穿到整个建站以及保护的进程之中。必定我们做网站不是为了给搜寻引擎看的而是给用户看的。站点得到用户的喜爱，那就是胜利的，所以将用户体验放在第一位才是 SEO 的核心宗旨。

1.8 SEO 的职业道德义务

网站排名所使用的 SEO 技巧可谓五花八门，各式各样。但不是所有的搜索引擎优化技术都是合乎道德规范的，这就引出一个关于搜索引擎优化的职业道德的问题。并且可以看到不少关于这方面的话题，但网上对此还未有一个较为明确的定义。我们听得最多的是所谓的“白帽子”和“黑帽子”，偶尔还会听到“灰帽子”这个词。在人们开始热衷于搜索引擎排名的今天，我们有必要想一想在这三大类别背后所蕴涵的真实意义。换而言之，职业道德究竟会不会对网站优化的具体实施策略起到直接的影响?

1.8.1 搜索引擎优化的职业道德

在探讨搜索引擎职业道德这个问题上，我们必须首先认识到这里面所涉及的问题是相当复杂的，而且带有很强的主观性。可能有人认为这种优化技术很好，但有人却觉得它拙劣不堪。所以，要想给搜索引擎优化技术用“优”、“良”和“差”进行划分的话，无异于是在瞄准一个飘乎不定的“移动靶”。而且最糟糕的是，由于我们对不同搜索引擎及各种不同的排名算法仍然缺乏全面深入的了解，因而要做这样一个性质上的准确划分也是几乎不可能的。但是我们可以给“白帽子”、“黑帽子”和“灰帽子”这三种技术下一个基本定义，这样读者就能够知道作者的态度：哪些技术合乎道德，哪些完全不符合道德，哪些又是介于两者之间的“灰色”技术。

1.8.2 搜索引擎优化之红、黑、灰帽

搜索引擎优化专业人员基本可以分为三大阵营：“白帽”SEO 人员，其对立阵营为“黑帽”人员，及介于两者之间的是“灰帽”人员。白帽优化技术其实就是已经得到大家广泛认可的优化技术，这些优化人员对 Google“网站管理指南”中所列出的种种被视为作弊的技术惟恐避之不及，同时杜绝使用一切带有任何可疑迹象的优化技术。

1. 白帽优化技术

标准的白帽优化技术包括：为网站添加更多主题相关的文字内容、定期更新网站的内容、为网站吸纳外部链接，并辅以合适的链接文字。保持网页标题的准确性和实时性。除此

之外，将重点放在网站内容的良好建设上，如增设 Blog 来丰富充实网站的内容，同时注意关键词在网站上的合理分布，一切工作都围绕让网站成为一个权威（Authority Site ）和中心网站（Hub Site）的目标来进行。

2. 黑帽优化技术

黑帽技术一般是指对网站使用在 Google 的“网站指南”中明令禁止的作弊优化方法，甚至包括被其他一些主流搜索引擎严禁的排名手段。如内容隐藏技术，重定向技术，链接工厂机制，以及大量交叉互链做法等，这些均属于黑帽技术，而这其中花样最多的要属隐藏内容技术。一般是把字体和背景颜色设为相同或几近相同的颜色，也有通过在 CSS 中将字体设为与背景相同或不可见，或通过层（Layer）来达到隐藏内容的目的。

3. 灰帽优化技术

灰帽技术介于这两者之间。这个概念比较难下定义，可以说亦正亦邪。一方面他们会使用一些白帽的优化技术，但同时又会用链接工具和内容生成软件类带有黑帽性质的优化技术。

1.8.3 搜索引擎优化的职业道德是否有意义

对于搜索引擎优化上的职业道德的探讨是否有意义？有些人会不假思索地说“有”；有些人的态度则比较犹疑；也有些人认为只要客户的排名能够上去，这一点根本无关紧要。总而言之，对这个问题并没有一个明确积极的答案。我们知道，一个搜索引擎优化人员要面对两个客户：这样他面对的就不止一种职业道德。首先，是他面对站长这个客户所需要具备的职业道德，其次是面对搜索引擎需要具备的职业道德。

但是，由于这两类客户在需求上存在冲突和矛盾，必然会导致搜索引擎优化人员陷入两头为难的困境。让网站获得尽可能高的排名是搜索引擎优化专业人员责无旁贷的工作，但同时也要照章办事。考虑到搜索引擎对网站的各种要求，优化人员必须知道如何划分网站和搜索引擎之间的道德界限。而且重要的是，他应就以下两个问题与客户进行认真讨论：一个是客户对网站排名的真正目的，其次就是客户对因优化而带来的风险的承受底线。比较专业的优化人员往往不会让客户涉足“黑帽子”甚至“灰帽子”这样的领域。此外，对于搜索引擎对网站的要求和网主对网站的排名要求之间的界限的理解，对各位优化专家来说则是仁者见仁，智者见智了。

1.8.4 优化与服务对象之间的关系

每个站长对自己网站未来的发展目标都不相同。有些网站要长足发展，所以自然希望他们的排名能够尽量保持持久。对这样一个希望具备一定知名度的网站，就需要用最优良和最地道的优化技术，从而使网站得以长期良好发展。有一些网站也希望他们的排名能够尽量保

持持久，但他们却很少留意过网站使用的优化技术是否得当。很多站点都游走在搜索引擎多种多样的 SPAM 过滤系统的边缘却不自知。事实上是，许多网站管理员在满怀喜悦地给网站做优化的时候，他们根本就没想到有些技术已经超出了搜索引擎的容忍范围。也有一些网站急于求成，根本不关心因优化手段的不正当性可能对网站带来的潜在风险。对这些网站来说，他们唯一的目的就是提高网站的排名，并为此穷尽一切可能手段，但对其正当性则置若罔闻。即使网站遭到搜索引擎的处罚或禁止，那也在他们的预料之内。

他们随之会再申请一个新的域名，然后故伎重演。许多搜索引擎优化人员都认为网站管理员这个客户是他们的第一服务对象，所以他们应该本着“一切以客户所需为中心”的理念，通过优化技术来满足客户的一切需求。不过，越是在这样的情况下，优化人员越有责任向客户提供详尽的信息，其中包括向客户说明哪些优化技术可能因为不被搜索引擎允许而可能给网站造成一定后果。如果不把这样的一些信息告知客户，其实对谁都没有好处。优化人员既然知道使用黑帽子或灰帽子优化技术可能会给网站造成一定负面影响（如网站可能会被搜索引擎处罚甚至删除，如果不把这样的信息告知客户，那就是对客户的不负责任，而且客户可能会在完全不知情的情况下来承担因为优化而给网站所造成的严重后果。此外，许多网站对搜索引擎优化知识都知之甚少，甚至有些人根本不知道还有这么一个行业存在）。在这种情况下，优化人员更应该把知情权还给客户，就优化的详细信息，包括所实施的优化技术可能让客户承担的风险通知客户，并让客户自行定夺。

1.8.5 优化与目标对象之间的关系

SEO 人员对搜索引擎的态度是另一个比较重要的因素。有些人把搜索引擎视作自己的朋友般友好，但有些优化人员却将其视做自己的死敌一样充满敌意。还有一些专业的优化人员要么将搜索引擎视作一个虽然不好但却不可缺少的东西，要么就把它当作一个整体互联网营销项目中的一个工具。对搜索引擎为什么会有这么多不同态度，原因其实并不难理解。

在那些将搜索引擎视做朋友的人眼中，搜索引擎所提供的搜索结果就意味着免费的客户源。搜索引擎优化要比其他网络推广方式经济得多，而且通过合理充分的优化，网站完全可以获得完全免费而且源源不断的客户源。他们深信：搜索引擎实际上是在向公众提供免费服务。正是基于这个原因，他们才将搜索引擎视做自己的朋友，并乐意为其提供经过精心打造的“最好”的网站。

有些优化人员却把搜索引擎几乎视为敌人。他们坚持认为：搜索引擎并没有向用户提供与查询最相关的搜索结果，反而让那些使用不正当优化手段的网站大行其道，而那些真正与用户查询最相关的高质量的网站却被排挤到角落里，结果是把搜索结果弄到一团糟。同时，也正是由于搜索结果缺乏足够相关性以及作弊网站充斥其间，抱有这种想法的优化人员认为搜索引擎根本就没有履行他们自称的目标或服务条款。同时，也正是由于搜索引擎的算法系统为作弊网站提供了可乘之机，因而“黑帽子”一方觉得，无论是他们的优化工作还是被优化的网站，都没有法律或道义上的责任以及义务来遵守搜索引擎制订的这些

规矩。他们认为：如果搜索结果让一些不怀好意的网站得逞，而规规矩矩遵守服务条款的网站却受到惩罚，那么根本没有理由再去遵守搜索引擎制订的所谓的“网站行为指南”了。

此外，还有一些优化人员把搜索引擎视作一种商业模式，而且搜索引擎的排名算法不过只是一种计算机程序，根本无所谓“好”和“坏”，它们所要执行的任务就是按网站对用户搜索条件的相关性来排列网站。而且正因为这是一种商业模式，如果要使用搜索引擎提供的业务或产品，用户需要遵守它们的商业规矩。在这种情况下，网站行为指南和服务条款就是搜索引擎为用户制订的商业规矩。就像饭店老板可以拒绝衣冠不整者入内就餐一样，搜索引擎也一样可以随时随地删除那些违反了它们的商业规矩的网站。

鉴于搜索引擎的职业道德及其排名算法上存在着太多争议，最好的办法就是把搜索引擎视做一种商业模式。把它们所制订的网站行为指南和服务条款视做商业规矩，对于不遵守这些规矩的网站，搜索引擎安全有理由从其数据库中自行删除。此外，搜索引擎也并无索引网络上的每个网页或站点这样的义务和职责。

对于那些不遵守它们规则的网站，它们完全有权予以惩罚或删除。所以如果不想让网站遭到搜索引擎的惩罚甚至删除，就需要网站遵守这些规则。至于那些对搜索引擎提供的搜索结果不满意的人，搜索引擎都设置了检举投诉体系，你完全可以向它们反映问题。在接到这样的投诉后，搜索引擎的客服中心一经核实一般都会给予删除。不过删除的过程比较漫长，有可能要等上几周甚至几个月。

虽然有人怀疑搜索引擎是否有能力改观对作弊网站的监测、惩罚以及删除，不过即使这个问题存在，它也并不意味着网站可以违反搜索引擎的服务条款而免受惩处。此外，一个作弊网站仍然占据搜索引擎。

对 SEO 职业道德所做的任何分析都要以站长对网站优化的真实意图为出发点。首先要考虑站长愿意承受的风险程度，即对被搜索引擎惩罚甚至遭封所持的态度。其次要考虑相关关键词的竞争性。最后一点非常重要，那就是需要考虑站长是不是会为达到排名而不择手段，无视一切职业道德。撇开道德这个因素不提，搜索引擎优化的最终目标无外乎让网站得到一个理想排名。让网站位居各大搜索引擎的显著位置实在是个巨大的诱惑，鲜有人面对这样的诱惑而无动于衷。

不管是为了增加网站的访问流量，提高销售额还是增加利润，或者仅仅只是为了满足自己的成就感，总之无论出于什么目的，网站管理员在优化策略上也是出尽百宝。这样就很可能面临这样一种状况，即你认为不错的优化技术却被别人所不耻。因而对一个网站所采用的优化技术是否合理（合乎道德），以及这些技术在搜索引擎优化中是否重要，就显得愈加重要了。举个例子，123 宠物网 http://www.123pet.com.cn 网站的首页，如图 1-5 所示。这个网站的优化主要是要站在爱宠一族的需求上去分析，满足需求、迎合搜索引擎、创造需求，才能更好地发展。



图 1-5　123 宠物网的首页

1.9 小结

搜索引擎优化虽然有很多的好处，而且各方面的效果都达到完美，但是 SEO 并不是万能的，它也有很多劣势，如工作时间较长、见效周期长、SEO 中关键字数量有限等。这就需要针对 SEO 的缺点，做一些其他方面的工作以达到利润最大化，效益最优化。

第 2 章　搜索引擎

说到搜索引擎大家都知道它是个效劳器，并且也了解搜索引擎是分为两个主体，一个是用来统计数据的搜索引擎效劳器，另一个则是收录页面信息的搜索引擎机器人，也就是蜘蛛，这两种缺一不可。本章主要为读者介绍搜索引擎的原理及工作流程。此外还将讲述搜索引擎各主要工作与搜索引擎优化之间的对应关系，为学习搜索引擎优化提供了理论依据，并且举出几种搜索引擎可能用到的排名计算方法。

2.1　搜索引擎的发展

在计算机网络的初始发展阶段，网站相对现在要少很多，信息查询也比较容易。然而伴随着互联网爆炸性的发展，普通网络用户想找到所需的资料简直如同大海捞针，这时为了满足大众信息检索需求的专业搜索网站便自然应运而生了。图 2-1 所示为成立较早的谷歌搜索引擎网站。



图 2-1　谷歌搜索引擎

现代意义上搜索引擎的祖先，是 1990 年由蒙特利尔大学学生 Alan Emtage 发明的 Archie。虽然当时 World Wide Web 还未出现，但网络中文件传输还是相当频繁的，而且由于大量的文件散布在各个分散的 FTP 主机中，查询起来非常不便，因此 Alan Emtage 想到了开发一个可以以文件名查找文件的系统，于是便有了 Archie。

Archie 工作原理与现在的搜索引擎很接近，它依靠脚本程序自动搜索网上的文件，然后对有关信息进行索引，供使用者以一定的表达式查询。由于 Archie 深受用户欢迎，受其启发，美国内华达 System Computing Services 大学于 1993 年开发了另一个与之非常相似的搜索工具，不过此时的搜索工具除了索引文件外还能检索网页。

当时，“机器人”一词在编程者中十分流行。电脑“机器人”（Computer Robot）是指某个能以人类无法达到的速度不间断地执行某项任务的软件程序。由于专门用于检索信息的机器人程序像蜘蛛一样在网络间爬来爬去，因此，搜索引擎的机器人程序就被称为蜘蛛程序。

世界上第一个用于监测互联网发展规模的机器人程序是 Matthew Gray 开发的 World wide Web Wanderer。刚开始它只用来统计互联网上的服务器数量，后来则发展为能够检索网站域名。

与 Wanderer 相对应，Martin Koster 于 1993 年 10 月创建了 ALIWEB，它是 Archie 的 HTTP 版本。ALIWEB 不使用机器人程序，而是靠网站主动提交信息来建立自己的链接索引，类似于现在的 Yahoo！。

随着互联网的迅速发展，使得检索所有新出现的网页变得越来越困难，因此，在 Matthew Gray 的 Wanderer 基础上，一些编程者将传统蜘蛛程序的工作原理做了些改进。其设想是，既然所有网页都可能有连向其他网站的链接，那么从跟踪一个网站的链接开始，就有可能检索整个互联网。到 1993 年底，一些基于此原理的搜索引擎开始纷纷涌现，其中以 JumpStation、The World Wide Web Worm（Goto 的前身，也就是今天的 Overture）和 Repository-Based Software Engineering（RBSE）spider 最负盛名。

然而 JumpStation 和 WWW Worm 只是以搜索工具在数据库中找到匹配信息的先后次序排列搜索结果，因此毫无信息关联度可言。而 RBSE 是第一个在搜索结果排列中引入关键字符串匹配程度概念的引擎。

现代意义上最早的搜索引擎出现于 1994 年 7 月。当时 Michael Mauldin 将 John Leavitt 的蜘蛛程序接入到其索引程序中，创建了大家现在熟知的 Lycos。同年 4 月，斯坦福（Stanford）大学的两名博士生，David Filo 和美籍华人杨致远（Gerry Yang）共同创办了超级目录索引 Yahoo！，并成功地使搜索引擎的概念深入人心。从此搜索引擎进入了高速发展时期。目前，互联网上有名有姓的搜索引擎已达数百家，其检索的信息量也与从前不可同日而语。如 Google，其数据库中存放的网页已达 40 亿之巨。

随着互联网规模的急剧膨胀，一家搜索引擎光靠自己单打独斗已无法适应目前的市场状况，因此现在搜索引擎之间开始出现了分工协作，并有了专业的搜索引擎技术和搜索数据库服务提供商。如国外的 Inktomi（已被 Yahoo！收购），它本身并不是直接面向用户的搜索引擎。而 Overture（原 GoTo，已被 Yahoo！收购）、LookSmart、MSN、HotBot 等搜索引擎提供全文网页搜索服务，国内的百度也属于这一类，搜狐和新浪用的就是它的技术。因此从这个意义上说，它们是搜索引擎的搜索引擎。

2.2 搜索引擎的分类

搜索引擎按其工作方式主要可分为三种，分别是全文搜索引擎（Full Text Search Engine）、目录索引类搜索引擎（Search Index/Directory）和元搜索引擎（Meta Search Engine）。

2.2.1 全文搜索引擎（Full Text Search Engine）

全文搜索引擎是名副其实的搜索引擎，国外代表有 Google，国内则有著名的百度搜索。它们从互联网提取各个网站的信息（以网页文字为主），建立起数据库，并能检索与用户查询条件相匹配的记录，按一定的排列顺序返回结果。

根据搜索结果来源的不同，全文搜索引擎可分为两类，一类拥有自己的检索程序（Indexer），俗称蜘蛛（Spider）程序或机器人（Robot）程序，能自建网页数据库，搜索结果直接从自身的数据库中调用，上面提到的 Google 和百度就属于此类；另一类则是租用其他搜索引擎的数据库，并按自定的格式排列搜索结果，如 Lycos 搜索引擎。

2.2.2 目录搜索引擎（Search Index/Directory）

目录搜索引擎是以人工方式或半自动方式搜集信息，由编辑员查看信息之后，人工形成信息摘要，并将信息置于事先确定的分类框架中。信息大多面向网站，提供目录浏览服务和直接检索服务。该类搜索引擎因为加入了人的智能，所以信息准确、导航质量高，缺点是需要人工介入、维护量大、信息量少、信息更新不及时。

目录索引虽然有搜索功能，但严格意义上不能称为真正的搜索引擎，只是按目录分类的网站链接列表而已。用户完全可以按照分类目录找到所需要的信息，不依靠关键词（Keywords）进行查询。目录索引中最具代表性的莫过于大名鼎鼎的 Yahoo！和新浪分类目录搜索，如图 2-2 所示。



图 2-2　雅虎搜索

首先，搜索引擎属于自动网站检索，而目录索引则完全依赖于手工操作。用户提交网站后，目录编辑人员会亲自浏览你的网站，然后根据一套自定的评判标准甚至编辑人员的主观印象，决定是否接纳你的网站。

其次，搜索引擎收录网站时，只要网站本身没有违反有关的规则，一般都能登录成功。而目录索引对网站的要求则高得多，有时即使登录多次也不一定成功。尤其像 Yahoo！这样的超级索引，登录更是困难。

此外，在登录搜索引擎时，一般不用考虑网站的分类问题，而登录目录索引时则必须将网站放在一个最合适的目录（Directory）中。

最后，搜索引擎中各网站的有关信息都是从用户网页中自动提取的，所以从用户的角度看，我们拥有更多的自主权；而目录索引则要求必须手工另外填写网站信息，而且还有各种各样的限制。更有甚者，如果工作人员认为你提交网站的目录、网站信息不合适，他可以随时对其进行调整，当然事先是不会和你商量的。

目录索引，顾名思义就是将网站分门别类地存放在相应的目录中，因此用户在查询信息时，可选择关键词搜索，也可按分类目录逐层查找。如以关键词搜索，返回的结果跟搜索引擎一样，也是根据信息关联程度排列网站，只不过其中人为因素要多一些。如果按分层目录查找，某一目录中网站的排名则是由标题字母的先后顺序决定的（也有例外）。

2.2.3 元搜索引擎（Meta Search Engine）

元搜索引擎（META Search Engine）接受用户查询请求后，同时在多个搜索引擎上搜索，并将结果返回给用户。著名的元搜索引擎有 InfoSpace、Dogpile、Vivisimo 等，中文元搜索引擎中具代表性的是搜星搜索引擎。在搜索结果排列方面，有的直接按来源排列搜索结果，如 Dogpile；有的则按自定的规则将结果重新排列组合，如 Vivisimo。

一个真正的元搜索引擎由三部分组成，即检索请求提交机制、检索接口代理机制、检索结果显示机制。

- 检索请求提交机制：负责实现用户“个性化”的检索设置要求，包括调用哪些搜索引擎、检索时间限制、结果数量限制等。
- 检索接口代理机制：负责将用户的检索请求“翻译”成满足不同搜索引擎“本地化”要求的格式。
- 检索结果显示机制：负责所有元搜索引擎检索结果的去重、合并、输出处理等。

元搜索引擎的出现，对于那些需要连续地使用不同的搜索引擎重复相同的检索的人来说，是一个福音。使用元搜索引擎同时对几个搜索引擎进行检索，获得分级编排的检索结果。

原理：我们可将元搜索引擎看成具有双层客户机/服务器结构的系统。用户向元搜索引擎发出检索请求，元搜索引擎再根据该请求向多个搜索引擎发出实际检索请求，搜索引擎执行元搜索引擎检索请求后将检索结果以应答形式传送给元搜索引擎，元搜索引擎将从多个搜索引擎获得的检索结果经过整理再以应答形式传送给实际用户。当然，某些元搜索引擎具有略微不同的机制，如图 2-3 所示。

分类：在可以检索的目标搜索引擎、检索提问的处理方式以及如何编译和显示结果方

面，元搜索引擎有着很大的差异。有些元搜索引擎一个接一个的搜索目标搜索引擎，另一些则同时进行搜索。有些搜索引擎将检索提问转变成目标搜索引擎的提问语言，而有一些则原封不动的发送给目标引擎。

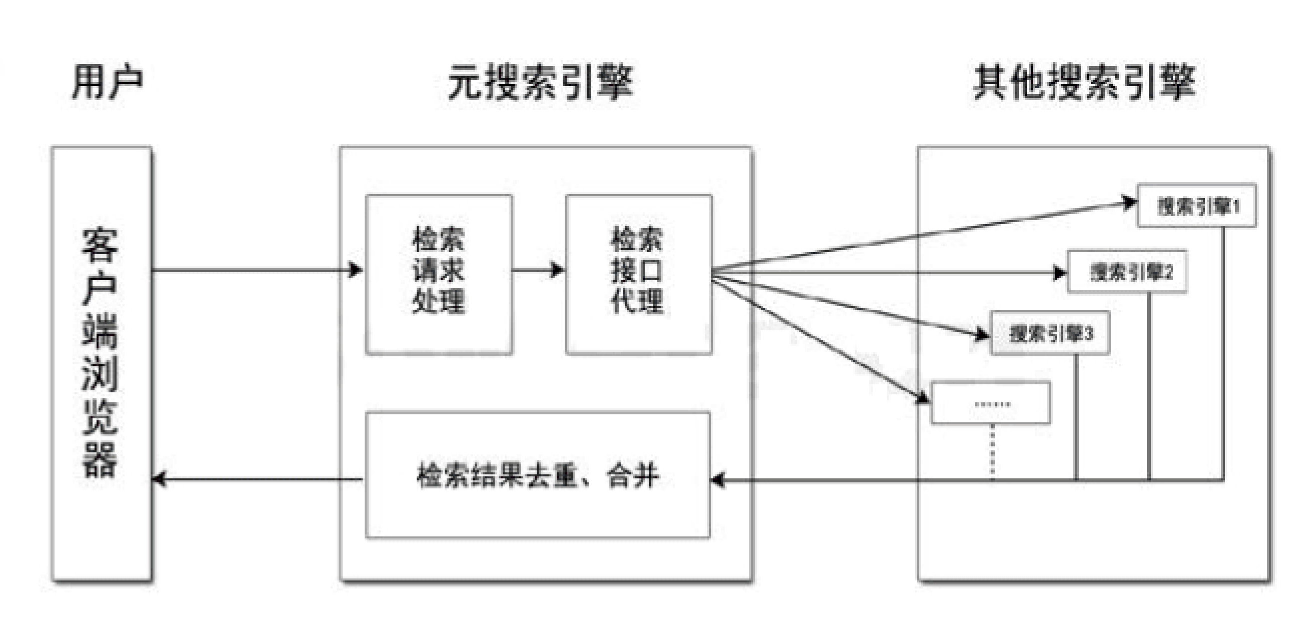


图 2-3　元搜索引擎工作原理

按功能划分，元搜索引擎包括多线索式搜索引擎和 All-in-One 式搜索引擎；按运行方式的差异可分为在线搜索引擎和桌面搜索引擎。

2.2.4　其他非主流搜索引擎形式

除以上介绍过的还有以下几种搜索引擎形式。

- 集合式搜索引擎：该搜索引擎类似元搜索引擎，区别在于它并非同时调用多个搜索引擎进行搜索，而是由用户从提供的若干搜索引擎中选择，如 HotBot 在 2002 年底推出的搜索引擎。
- 门户搜索引擎：AOL Search、MSN Search 等虽然提供搜索服务，但自身既没有分类目录也没有网页数据库，其搜索结果完全来自其他搜索引擎。
- 免费链接列表（Free For All Links，简称 FFA）：一般只简单地滚动链接条目，少部分有简单的分类目录，不过规模要比 Yahoo! 等目录索引小很多。

2.3　搜索引擎的工作流程、原理及作用

搜索引擎的工作流程对于外界来说神秘而高不可测，其实搜索引擎的工作流程非常简单。当作为一般用户在使用搜索引擎的时候，很难去想象或者了解搜索引擎的工作原理是怎么样的？所以正是因为对搜索引擎的工作流程不是很了解，导致现在很多人都认为搜索引擎返回的结果是动态的，也就是当用户输入一个语句进行关键字查询的时候，搜索引擎马上从互联网上抓取筛选结果，然后把查询的关键字结果反馈给用户。其实搜索引擎返回的结果是提前就已经抓取好了的，然后经过了一系列算法筛选之后放入数据库。当用户查询时就立即

对索引数据库进行查找，然后反馈给用户准确的关键字查询结果。

2.3.1 搜集信息形成快照

网页快照收录是搜索引擎通过蜘蛛程序在互联网上抓取页面并进行存储形成快照的过程，他为搜索引擎开展各项工作提供了数据支持，如图 2-4 所示。



图 2-4 123 宠物网百度快照

搜索引擎的信息搜集基本都是自动的。搜索引擎利用称为网络蜘蛛（spider）的自动搜索机器人程序来连上每一个网页上的超链接。机器人程序根据网页链到其他网页中的超链接，就像日常生活中所说的“一传十，十传百……”一样，从少数几个网页开始，连到数据库上所有到其他网页的链接。理论上，若网页上有适当的超链接，机器人便可以遍历绝大部分网页。

2.3.2 整理信息分析页面

普通用户看页面，看到的都是文字、图片等直观信息，但是搜索引擎看的是网页的源代码，而且会根据这些源代码来确定网页的重要信息。

在看一个页面的时候，会发现一些粗体、斜体、加色等样式的文字，还有一些比较大的字体，有些文字还像标题一样被隔开，这些设置不光是让用户在浏览信息的时候要特别注意，还有一点就是它们可能是网页的重点。

在搜索引擎看来，源代码头部或者靠近头部的信息往往是比其他位置的文字重要，大家可以很形象的将一个网页和一张报纸的页面比较着看。确定重要文字或词语，这是搜索引擎建立网页索引过程的一个步骤。

搜索引擎整理信息的过程称为“建立索引”。搜索引擎不仅要保存搜集起来的信息，还要将它们按照一定的规则进行编排。这样，搜索引擎根本不用重新翻查它所有保存的信息就能迅速找到用户所需要的资料。想象一下，如果信息是不按任何规则地随意堆放在搜索引擎的数据库中，那么它每次找资料都得把整个资料库完全翻查一遍，如此一来，再快的计算机系统也没有用。

2.3.3 接受查询处理关键词

搜索引擎还会根据同义词典和潜在语义进行拓展查询检索项。如搜索引擎会将“宠物”

和“狗狗”，“藏獒”和“犬”，“人民”和“百姓”当做同一个概念进行处理。另外，搜索引擎在查询的时候，还会根据主题来进行归类，如当查询 SEO 的时候，系统会自动将“网络营销”、“搜索引擎优化”、“网站优化”之类的都归为一个主题来处理。

当用户使用搜索引擎进行查询的时候，在搜索框中输入一个关键词，当索引中有这样一个完全匹配的结果的时候，就会显示出来。对于英文来说，相对就比较简单了，只要对英文的单复数，时态、词根、组合词等进行分析即可。不过对于中文，则一定要进行分词，然后对关键词进行向量计算，得出哪个网站的哪个关键词的权重比较高，就显示在索引中。

用户向搜索引擎发出查询，搜索引擎接受查询并向用户返回资料。搜索引擎每时每刻都要接到来自大量用户的几乎是同时发出的查询，它按照每个用户的要求检查自己的索引，在极短时间内找到用户需要的资料，并返回给用户。目前，搜索引擎返回主要是以网页链接的形式提供的，通过这些链接，用户便能到达含有自己所需资料的网页。通常搜索引擎会在链接下提供一小段来自这些网页的摘要信息以帮助用户判断此网页是否含有自己需要的内容。

2.4 搜索引擎对网页的收录

2.4.1 页面收录的流程

在互联网中，URL 是每个页面的入口地址，蜘蛛程序就是通过这些 URL 列表抓取到页面的，蜘蛛不断的从这些页面中获取 URL 资源及存储页面，并加入 URL 列表，如此不断的循环，搜索引擎就可以从互联网中获取到足够的页面，如图 2-5 所示。

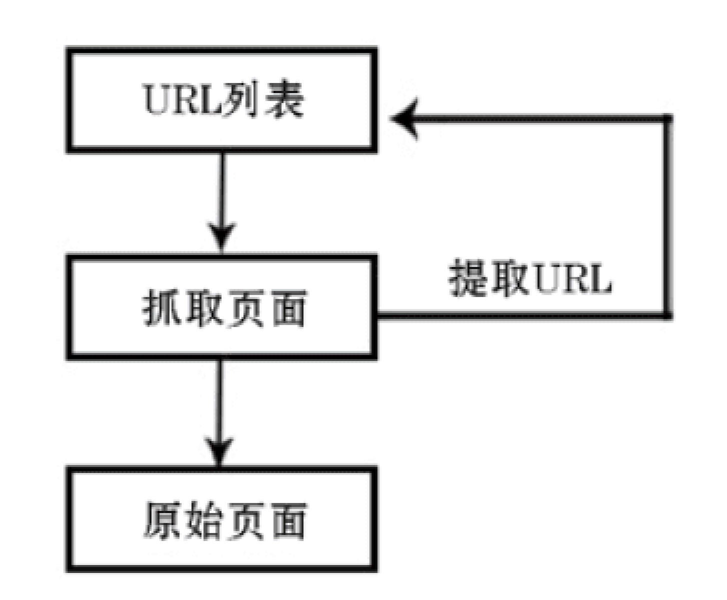


图 2-5　搜索引擎抓取页面简单流程

URL 是页面的入口，域名则是网站的入口。搜索引擎就是通过域名进入网站，挖掘 URL 资源。换而言之，搜索引擎在互联网中抓取页面的首要任务就是要有庞大的域名列表，再不断地通过域名，进入网站抓取网站中的页面。

而对于 SEO 而言，想被搜索引擎收录，首要条件就是加入搜索引擎的域名列表，常见加入搜索引擎的域名列表的有以下两种方式。

（1）利用搜索引擎提供的网站登录入口，向搜索引擎提交网站域名，如百度的：http://www.baidu.com/search/url_submit.html，可在此提交自己的网站域名，不过用此方法搜索引擎只会定期进行抓取并更新，这种做法比较被动，从域名提交到网站被收录所花费的时间也比较长。

（2）通过与有质量的“外链”，使搜索引擎在抓取“别人”的网站页面时发现我们的网站，从而实现对网站的收录，这种方法主动权在我们手上（只要有足够多的“外链”）且收录速度比第一种方法快，根据外部链接的数量、质量相关性，一般 2~7 天就会被搜索引擎收录。

2.4.2 页面收录的原理

通过前面的介绍，读者大致了解了搜索引擎的工作方式，下面我们来介绍页面收录的原理。页面收录原理经过修改“页面收录流程”能够把握放慢网站被收录的办法，从而改善搜索引擎收录的数目。假设把一个网站页面组成的页面看作是一个有向图，沿着页面中的链接，依照某种特定的战略对网站中的页面停止遍历。不停地从 URL 列表中移出曾经拜访的 URL，同时提取原始页面中的 URL 的消息，再将 URL 分为域名及外部 URL 两大类，同时判定 URL 能否被拜访过，递归地扫描 URL 列表，直至耗尽一切 URL 资源为止，如图 2-6 所示。

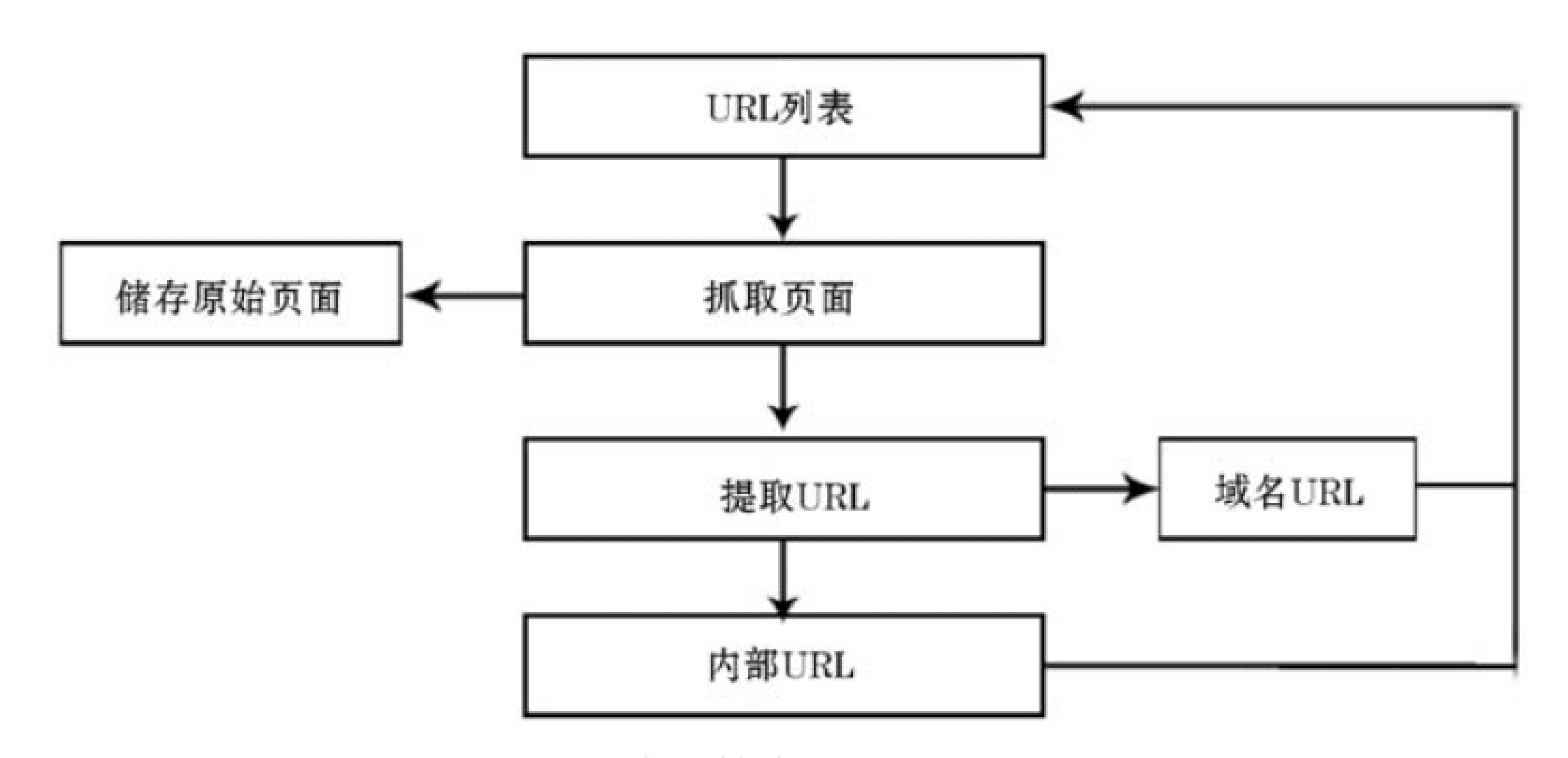


图 2-6　搜索引擎收录页面的工作原理

在搜索引擎中要获取相对重要的页面，就涉及到了搜索引擎的页面收录方式。页面收录方式是指搜索引擎抓取页面时所使用的战略，目的是为了能在互联网中挑选出绝对主要的消息。假设使用相同的抓取战略，搜索引擎在异样的工夫内能够在某一网站中抓取到更多的页面资源，则会在该网站停止更长的工夫，收录的页面数自然也就多了。因而，加强对搜索引擎页面收录方式的熟悉，有益于为网站树立敌对的构造，进步被收录的数目。

搜索引擎收录页面的方式主要有“广度优先”、“深度优先”及“用户提交”（用户提交暂时不讲）三种。

1. 广度优先

如果把整个网站看做一棵树，首页就是根，每个页面就是叶子。广度优先是一种横向的页面抓取方式，先从树的较浅层开始抓取页面，直接抓完同层次的所有页面后才进入下一层。因此，在对网站进行优化时，应该把网站相对重要的信息展示在层次比较浅的页面上（如在首页推荐一些热门的内容）。反过来，通过广度优先的抓取方式，搜索引擎就可以首先抓取到网站中相对重要的页面。

首先，蜘蛛从网站的首页出发，抓取首页上所有连接指向的页面，形成页面集合 A，并分析出 A 中所有页面中的链接，再跟踪这些链接抓取下一层的页面，形成页面集合 B。就这样递归地从浅层页面中解析出链接，再延伸到深层页面，直到满足某个设定的条件才停止抓取进程，如图 2-7 所示。

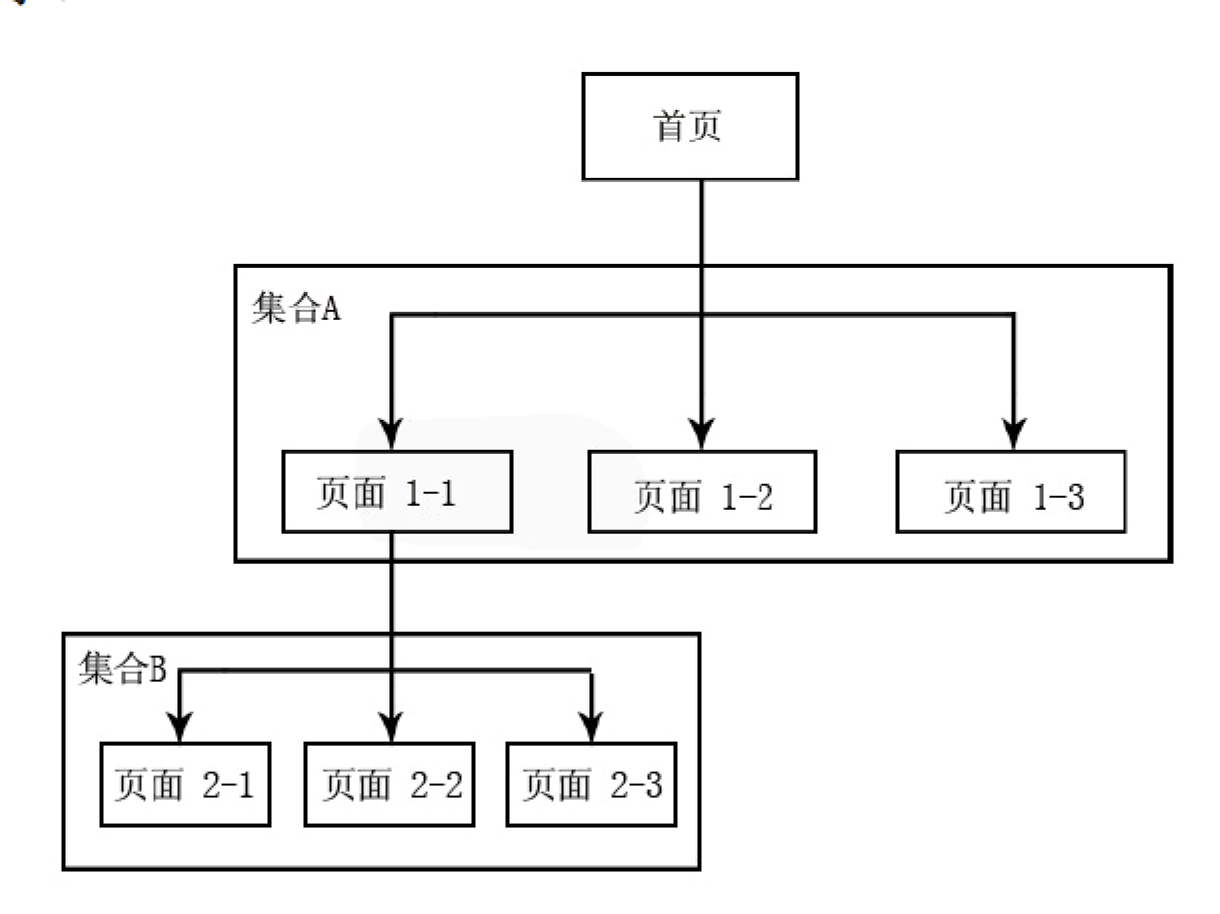


图 2-7　广度优先抓取流程

2. 深度优先

与广度优先的抓取方式相反，深度优先首先跟踪浅层页面中的某一连接后逐步抓取深层页面，直到抓完最深层的页面才返回浅层页面再跟踪另一链接，继续向深层页面抓取，这是一种纵向的页面抓取方式。使用深度优先的抓取方式，搜索引擎可以抓取到网站中较为隐蔽、冷门的页面，这样就能满足更多用户的需求。

首先，搜索引擎会抓取网站的首页，并提取首页中的链接，再沿着其中的一个连接抓取到页面 A-1，同时获取 A-1 中的链接并抓取页面 B-1，获取 B-1 中的来链接并抓取页面 C-1，如此不断地重复，满足到某个条件后，再从 A-2 抓取页面及链接，如图 2-8 所示。

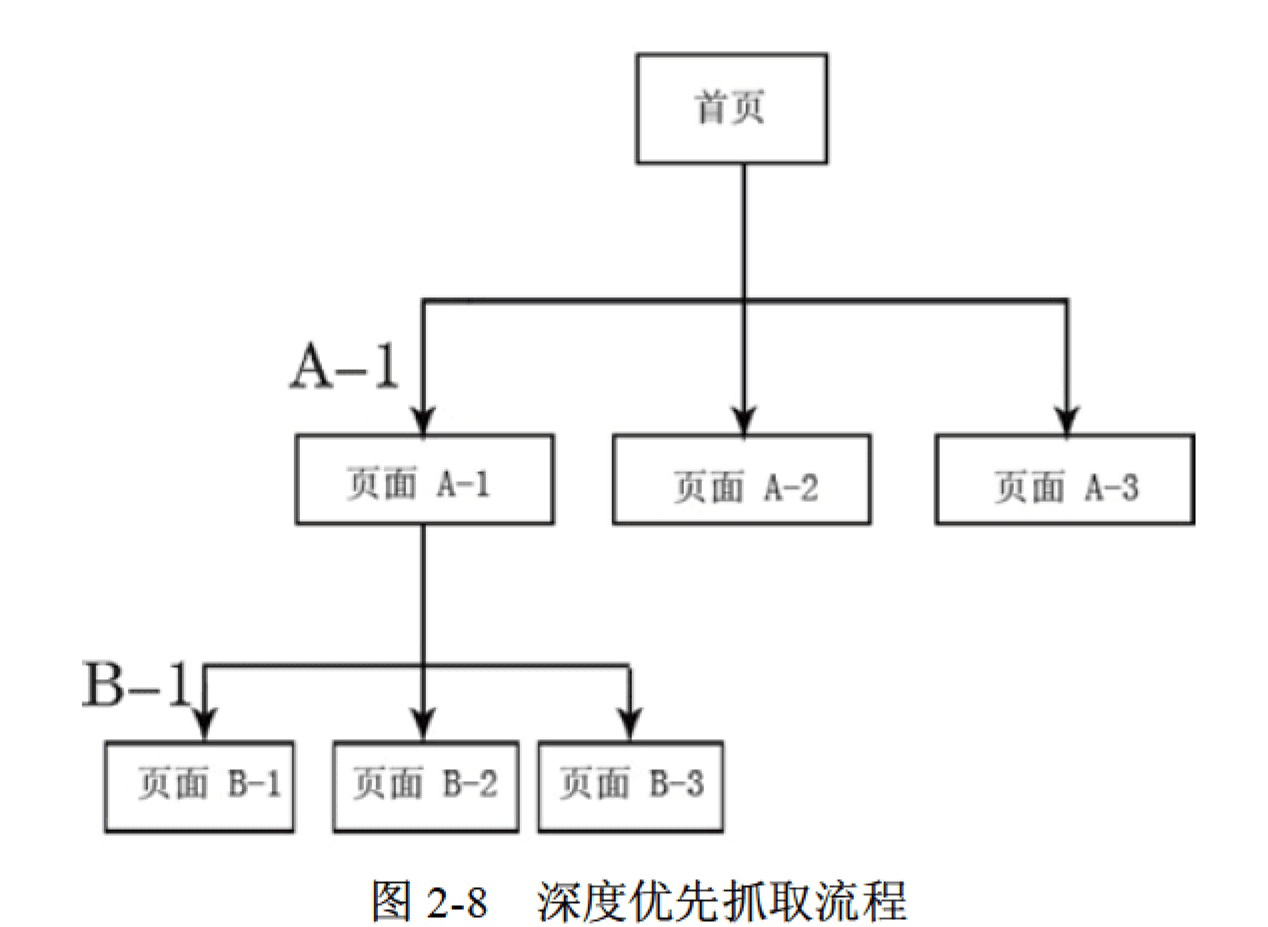


图 2-8　深度优先抓取流程

2.4.3　搜索引擎页面排名原理

在 Google 诞生以前，传统搜索引擎主要依靠页面内容中的关键词匹配搜索词进行排名。这种排名方式的短处现在看来显而易见，那就是很容易被刻意操纵。黑帽 SEO 在页面上堆积关键词，或加入与主题无关的热门关键词，都能提高排名，使搜索引擎排名结果质量大为下降。现在的搜索引擎都使用链接分析技术、减少垃圾，提高用户体验。本节将简要介绍链接在搜索引擎排名中的应用原理。

在排名中计入链接因素，不仅有助于减少垃圾，提高结果相关性，也使传统关键词匹配无法排名的文件能够被处理。如图片、视频文件无法进行关键词匹配，但是却可能有外部链接，通过链接信息，搜索引擎就可以了解图片和视频的内容从而排名。

不同文字的页面排名也成为可能。如在百度或 google.cn 搜索 SEO，都可以看到英文和其他文字的 SEO 网站。甚至搜索“搜索引擎优化”，也可以看到非中文页面，原因就在于有的链接可能使用“搜索引擎优化”为锚文字指向英文页面。

链接因素现在已经超过页面内容的重要性。链接关系比较抽象，不容易理解。但页面上的因素对排名的影响能看得到，容易直观理解。举个简单的例子，搜索一个特定关键词，SEO 人员只要观察前几页结果，就能看到。关键词在标题标签中出现有什么影响，出现在最前面又有什么影响，有技术资源的还可以大规模地统计，计算出关键词出现在标题标签中不同位置与排名之间的关系。虽然这种关系不一定是因果关系，但至少是统计上的联系，使 SEO 人员大致了解该如何优化。

链接对排名的影响就无法直观了解了，也很难进行统计，因为没有人能获得搜索引擎的链接数据库。我们能做的只是定性观察和分析。

下面介绍一些关于链接的专利，从而了解链接在搜索引擎排名中的使用方法和地位。

1. 李彦宏超链分析专利

百度创始人李彦宏在回国创建百度之前就是美国顶级的搜索引擎工程师之一。据说李彦

宏在寻找风险投资时，投资人询问其他三个搜索引擎业界的技术高人一个问题：要了解搜索引擎技术应该问谁。这三个被问到的高人中有两个回答：搜索引擎的事就问李彦宏。由此投资人断定李彦宏是最了解搜索引擎的人之一。

这其实就是现实生活中类似于链接关系的应用。要判断哪个页面最具权威性，不能光看页面自己怎么说，还要看其他页面怎么评价。

李彦宏 1997 年就提交了一份名为“超链文件检索系统和方法”的专利申请，这比 Google 创始人发明 PR 要早得多，不得不说这是非常具有前瞻性的研究工作。在这份专利中，李彦宏提出了与传统信息检索系统不同的基于链接的排名方法。

这个系统除了索引页面之外，还建立一个链接词库，记录链接锚文字的相关信息，如锚文字中包含哪些关键词，发出链接的页面索引，包含特定锚文字的链接总数，包含特定关键词的链接都指向哪些页面。词库不仅包含关键词原型，也包含同一个词的其他衍生关键词。

根据这些链接数据，尤其是锚文字，计算出基于链接的文件相关性。在用户搜索时，将得到的基于链接的相关性与基于关键词匹配的传统相关性综合使用，得到更准确的排名。

在今天看来，这种基于链接的相关性计算是搜索引擎的常态，每个 SEO 人员都知道。但是在十三四年前，这无疑是非常创新的概念。当然现在的搜索引擎算法对链接的考虑，已经不仅仅是锚文字，实际上要复杂得多。

2. HITS 算法

HITS 是英文 Hyperlink-Induced Topic Search 的缩写，意译为“超链诱导主题搜索”。HITS 算法由 Jon Kleinberg 于 1997 年提出，并申请了专利。

按照 HITS 算法，用户输入关键词后，算法对返回的匹配页面计算两种值，一种是枢纽值（Hub Scores），另一种是权威值（Authority Scores），这两种值是相互依存、相互影响。所谓枢纽值，是指页面上所有导出链接指向页面的权威值之和。权威值是指所有导入链接所在页面的枢纽值之和。

上面的定义比较拗口，简单地说，HITS 算法会提炼出两种比较重要的页面，也就是枢纽页面和权威页面。枢纽页面本身可能没有多少导入链接，但是有很多导出链接指向权威页面。权威页面本身可能导出链接不多，但是有很多来自枢纽页面的导入链接。

典型的枢纽页面有雅虎目录、开放目录或好 123 等网站目录。这种高质量的网站目录作用就在于指向其他权威网站，所以称为枢纽。而权威页面有很多导入链接，其中包含很多来自枢纽页面的链接。权威页面通常是提供真正相关内容的页面。

HITS 算法是针对特定的查询词，所以称为主题搜索。

HITS 算法的最大缺点是，它在查询阶段进行计算，而不是在抓取或预处理阶段。所以 HITS 算法是以牺牲查询排名响应时间为代价的。也正因为如此，原始 HITS 算法在搜索引擎中并不常用。不过 HITS 算法的思想很可能融入到搜索引擎的索引阶段，也就是根据链接关系找出具有枢纽特征或权威特征的页面。

成为权威页面是第一优先的，不过难度比较大，唯一的方法就是获得高质量链接。当你的网站不能成为权威页面时，就让它成为枢纽页面。所以导出链接也是当前搜索引擎排名因

素之一。绝不链接到其他网站的做法，并不是好的 SEO 方法。

3. TrustRank 算法

TrustRank 是近年来比较受关注的基于链接关系的排名算法。TrustRank 可以翻译为“信任指数”。

TrustRank 算法最初来自于 2004 年斯坦福大学和雅虎的一项联合研究，用来检测垃圾网站，并且于 2006 年申请专利。TrustRank 算法发明人还发表了一份专门的 PDF 文件，说明 TrustRank 算法的应用。

TrustRank 算法并不是由 Google 提出的，不过由于 Google 所占市场份额最大，而且 TrustRank 在 Google 排名中也是一个非常重要的因素，所以有些人误以为 TrustRank 是 Google 提出的。更让人糊涂的是，Google 曾经把 TrustRank 申请为商标，但是 TrustRank 商标中的 TrustRank 指的是 Google 检测含有恶意代码网站的方法，而不是指排名算法中的信任指数。

TrustRank 算法基于一个基本假设：好的网站很少会链接到坏的网站。反之则不成立，也就是说，坏的网站很少链接到好网站这句话并不成立。正相反，很多垃圾网站会链接到高权威、高信任指数的网站，试图提高自己的信任指数。

基于这个假设，如果能挑选出可以百分之百信任的网站，这些网站的 TrustRank 评为最高，这些 TrustRank 最高的网站所链接到的网站信任指数稍微降低，但也会很高。与此类似，第二层被信任的网站链接出去的第三层网站，信任度继续下降，由于种种原因，好的网站也不可避免地会链接到一些垃圾网站。不过离第一层网站点击距离越近，所传递的信任指数越高，离第一级网站点击距离就越远，信任指数将依次下降。这样，通过 TrustRank 算法，就能给所有网站计算出相应的信任指数，离第一层网站越远，成为垃圾网站的可能性就越大。

计算 TrustRank 值首先要选择一批种子网站，然后人工查看网站，设定一个初始 TrustRank 值。挑选种子网站有两种方式，一种是选择导出链接最多的网站，因为 TrustRank 算法就是计算指数随着导出链接衰减。导出链接多的网站，在某种意义上可以理解为“逆向 PR 值”比较高。

另一种挑选种子网站的方法是选 PR 值高的网站，因为 PR 值越高，在搜索结果页面出现的概率就越大。这些网站才正是 TrustRank 算法最关注的、需要调整排名的网站。那些 PR 值很低的页面，在没有 TrustRank 算法时排名也很靠后，计算 TrustRank 意义就不大了。

根据测算，挑选出大概两百个网站作为种子，就可以比较精确地计算出所有网站的 TrustRank 值。

计算 TrustRank 随链接关系减少的公式有两种方式。一种是随链接次数衰减，也就是说如果第一层页面 TrustRank 指数是 100，第二层页面衰减为 90，第三层衰减为 80。第二种计算方法是按导出链接数目分配 TrustRank 值，也就是说，如果一个页面的 TrustRank 值是 100 页面上有 5 个导出链接，每个链接将传递 20%的 TrustRank 值。衰减和分配这两种计算方法通常综合使用，整体效果都是随着链接层次的增加，TrustRank 值逐步降低。

得出网站和页面的 TrustRank 值后，可以通过两种方式影响排名。一种是把传统排名算法挑选出的多个页面，根据 TrustRank 值比较，重新做排名调整。另一种是设定一个最低的 TrustRank 值门槛，只有超过这个门槛的页面，才被认为有足够的质量进入排名，低于门槛的页面将被认为是垃圾页面，从搜索结果中过滤出去。

虽然 TrustRank 算法最初是作为检测垃圾的方法，但在现在的搜索引擎排名算法中，TrustRank 概念使用更为广泛，常常影响大部分网站的整体排名。TrustRank 算法最初针对的是页面级别，现在在搜索引擎算法中，TrustRank 值也通常表现在域名级别，整个域名的信任指数越高，整体排名能力就越强。

4. Google PR

PR 是 PageRank 的缩写。Google PR 理论是所有基于链接的搜索引擎理论中最有名的。SEO 人员可能不清楚本节介绍的其他链接理论，但不可能不知道 PR。

PR 是 Google 创始人之一拉里佩奇发明的，用于表示页面重要性的概念。用最简单的话说就是，反向链接越多的页面就是越重要的页面，因此 PR 值也就越高。

Google PR 有点类似于科技文献中互相引用的概念，被其他文献引用较多的文献，很可能是比较重要的文献。

（1）PR 的概念和计算

把互联网理解为由结点及链接组成的有向图，页面就是一个个结点，页面之间的有向链接传递着页面的重要性。一个链接传递的 PR 值决定于导入链接所在页面的 PR 值，发出链接的页面本身 PR 值越高，所能传递出去的 PR 也越高。传递的 PR 数值也取决于页面上的导出链接数目。对于给定 PR 值的页面来说，假设能传递到下级页面 100 份 PR，页面上有 10 个导出链接，每个链接能传递 10 份 PR，页面上有 20 个导出链接的话，每个链接只能传递 5 份 PR。所以一个页面的 PR 值取决于导入链接总数，发出链接页面的 PR 值，以及发出链接页面上的导出链接数目。

PR 值计算公式如下：

```
PR(A)=(1-d)+d(PR(t1)/C(t1)+ … +PR(tn)/C(tn))
```

A 代表页面 A。

PR（A）则代表页面 A 的 PR 值。

d 为阻尼指数。通常认为 d=0.85。

t1...tn 代表链接向页面 A 的页面从 t1 到 tn。

C 代表页面上的导出链接数目。C（t1）即为页面 t1 上的导出链接数目。

从概念及计算公式都可以看到，计算 PR 值必须使用迭代计算。页面 A 的 PR 值取决于链接向 A 的页面 t1~tn 页面的 PR 值，而 t1~tn 页面的 PR 值又取决于其他页面的 PR 值，其中很可能还包含页面 A。所以 PR 需要多次迭代才能得到。计算时先给所有页面设定一个初始

值，经过一定次数的迭代计算后，各个页面的 PR 值将趋于稳定。研究证明，无论初始值怎么选取，经过迭代计算的最终 PR 值都不会受到影响。

外部页面 Y 向循环注入 PR 值，循环中的页面不停迭代传递 PR，如果没有阻尼系数，循环中的页面 PR 将达到无穷大。引入阻尼系数，使 PR 在传递时自然衰减，才能将 PR 计算稳定在一个值上。

（2）PR 的两个比喻模型

关于 PR 有两个著名的比喻。

- 一个比喻是投票。链接就像民主投票一样，A 页面链接到 B 页面，就意味着 A 页面对 B 页面投了一票，使得 B 页面的重要性提高。同时，A 页面本身的 PR 值决定了 A 所能投出去的投票力，PR 值越高的页面，投出的票也更重要。在这个意义上，传统基于关键词匹配的算法是看页面自己说页面内容是什么，基于链接的 PR 则是看别人怎么评价一个页面。
- 另一个比喻是随机冲浪。假设一个访问者从一个页面开始，不停地随机点击链接，访问下一个页面。有时候这个用户感到无聊了，不再点击链接，就随机跳到了另外一个网址，再次开始不停地向下点击。所谓 PR 值，也就是一个页面在这种随机冲浪访问中被访问到的概率。一个页面导入链接越多，被访问到的概率也越高，因此 PR 值也越高。

阻尼系数也与随机冲浪模型有关。（1-d）=0.15 实际上就是用户感到无聊，停止点击，随机跳到新 URL 的概率。

5. Hilltop 算法

Hilltop 算法是由 Krishna Baharat 在 2000 年左右所研究的，于 2001 年申请了专利，并且把专利授权给 Google 使用，后来 Krishna Baharat 本人也加入了 Google。

Hilltop 算法可以简单地理解为与主题相关的 PR 值。传统 PR 值与特定关键词或主题没有关联，只计算链接关系。这就有可能出现某种漏洞。如一个 PR 值极高的关于环保内容的大学页面，上面有一个链接连向一个儿童用品网站，这个链接出现的原因可能仅仅是因为这个大学页面维护人是个教授，他的太太在那个卖儿童用品的公司工作。这种与主题无关、却有着极高 PR 值的链接，有可能使一些网站获得很好的排名，但其实相关性并不高。

Hilltop 算法就尝试矫正这种可能出现的疏漏。Hilltop 算法同样是计算链接关系，不过它更关注来自主题相关页面的链接权重。在 Hilltop 算法中把这种主题相关页面称为专家文件。显然，针对不同主题或搜索词有不同的专家文件。

根据 Hilltop 算法，用户搜索关键词后，Google 先按正常排名算法找到一系列相关页面并排名，然后计算这些页面有多少来自专家文件的、与主题相关的链接，来自专家文件的链接越多，页面的排名分值越高。按 Hilltop 算法的最初构想，一个页面至少要有两个来自专家文件的链接，才能返回一定的 Hilltop 值，不然返回的 Hilltop 值将为零。

根据专家文件链接计算的分值被称为 LocalRank。排名程序根据 LocalRank 值，对原本传统排名算法计算的排名做重新调整，给出最后排名。这就是前面介绍的搜索引擎排名阶段最后的过滤和调整步骤。

Hilltop 算法最初写论文和申请专利时对专家文件的选择有不同描述。在最初的研究中，Krishna Baharat 把专家文件定义为包含特定主题内容，并且有较多的导出链接到第三方网站的页面，这有点类似于 HITS 算法中的枢纽页面。专家文件链接指向的页面与专家文件本身没有关联，但是它们来自同一个主域名下的子域名，来自相同或相似 IP 地址的页面等。最常见的专家文件经常来自于学校、政府及行业组织网站。

在最初的 Hilltop 算法中，专家文件是预先挑选的。搜索引擎可以根据最常见的搜索词，预先计算出一套专家文件，用户搜索时，排名算法从事先计算的专家文件集合中选出与搜索词相关的专家文件子集，再从这个子集中的链接计算 LocalRank 值。

不过在 2001 年所申请的专利中，Krishna Baharat 描述了另外一个挑选专家文件的方法，专家文件并不预先选择，用户搜索特定查询词后，搜索引擎按传统算法挑出一系列初始相关页面，这些页面就是专家文件。Hilltop 算法在这个页面集合中再次计算哪些网页有来自于集合中其他页面的链接，赋予比较高的 LocalRank 值。由于传统算法得到的页面集合已经具备了相关性，这些页面再提供链接给某一个特定页面，这些链接的权重自然应该很高。这种挑选专家文件的方法是实时进行的。

通常认为 Hilltop 算法对 2003 年底的佛罗里达更新有重大影响，不过 Hilltop 算法是否真的已经被融入进 Google 排名算法中，没有人能够确定。Google 从来没有承认、也没有否认自己的排名算法中是否使用了某项专利。不过从排名结果观察及招揽 Krishna Baharat 至麾下等迹象看，Hilltop 算法的思想得到了 Google 的极大重视。

Hilltop 算法提示 SEO，建设外部链接时更应该关注主题相关的网站。最简单的方法是搜索某个关键词，目前排在前面的页面就是最好的链接来源，甚至可能有一个来自竞争对手网站的链接，这种效果是最好的。当然，获得这样的链接难度最大。

2.4.4 搜索引擎模块组成

一个典型的网络信息检索系统的架构由信息收集、查询服务和信息处理三个模块组成。

从具体运行方式上说，系统根据站点，网页的 URL 信息和网页之间的链接关系，利用网络蜘蛛在互联网上收集数据。收集的数据分别通过链接信息分析器和文本信息分析器处理，保存在链接数据库和文本索引数据库中，同时，网页质量评估器依据网页的链接关系和页面结构特征对页面质量进行评估，并将评估的结果保存在索引数据库中。查询服务器负责与用户的交互，它根据用户的检索需求，从索引数据库中读取对应的索引，并综合考虑查询相关性与页面质量评估结果之间的关系，给出查询结果列表反馈给用户。

1. 信息收集模块

搜索器的信息收集模块包括“蜘蛛控制”和“网络蜘蛛”两部分，“蜘蛛”这个称呼形

象的描述出了信息收集模块在网络数据形成的 Web 上进行信息获取的功能。总体而言，网络蜘蛛从种子网页出发，通过反复下载网页并从文档中寻找未曾见过的 URL，达到访问其他网页得以遍历 Web 的目的。而其工作策略一般则可以分为累积式抓取（cumulative crawling）和增量式抓取（incremental crawling）两种。

累积式抓取是指从某一个时间点开始，通过遍历的方式抓取系统所能允许存储和处理的所有网页。在理想的软硬件环境下，经过足够的运行时间，可以抓取到相当规模的网页集合。但由于 Web 数据的动态特性，集合中网页的被抓取时间点是不同的，页面被更新的情况也不同，因此累积式抓取到的网页集合事实上无法与真实环境中的网络数据保持一致。

与累积式抓取不同，增量式抓取是指在具有一定量规模的网络页面集合的基础上，采用更新数据的方式对已有集合中的过时网页进行抓取，以保证所抓取到的数据与真实网络数据足够接近。进行增量式抓取的前提是，系统已经抓取了足够数量的网络页面，并具有这些页面被抓取的时间信息。

面向实际应用环境的网络蜘蛛设计中，通常既包括累积式抓取，也包括增量式抓取的策略。累积式抓取一般用于数据集合的整体建立或大规模更新阶段；而增量式抓取则主要针对数据集合的日常维护与即时更新。

在确定了抓取策略之后，如何充分利用网络带宽，并合理确定网页数据更新的时间点就成了网络蜘蛛运行策略中的核心问题。

总体而言，在合理利用软硬件资源进行针对网络数据的即时抓取方面，已经形成了比较成熟的技术和实用性解决方案，这方面目前所需解决的主要问题，是如何更好地处理动态网络数据问题（如数量越来越庞大的 Web 2.0 数据等），以及更好地根据网页质量修正抓取策略的问题。

2. 查询服务模块

查询服务模块是指网络信息检索系统中，直接与用户查询需求交互的组成部分。从功能上讲，这个模块主要完成两方面的任务：首先是用户交互，即收集用户查询并用符合用户行为习惯的方式对查询结果进行反馈；其次是信息查询和检索，使用收集到的用户查询在由信息收集模块抓取，并由信息处理模块建立索引的网络数据进行检索。这其中涉及到的核心技术包括：检索系统网络服务的建立和维护、面向网络信息的检索模型建立等。

网络信息检索系统严格意义上说也在提供一种网络服务（Web Service），因此它的建立和运行、维护机制符合网络服务系统的一般规律。但由于网络信息检索用户行为习惯在宏观上的统计特性，系统的网络服务也应进行优化以更好地适应这种行为习惯。针对搜索引擎日志分析的工作说明，尽管网络信息检索用户所提出的查询数目异常繁杂，但搜索需求中的大多数是重复的，重复的比例超过用户所提交查询需求的 70%以上乃至更多。因此在检索系统的网络服务部分加入某种类型的缓存机制就是非常必要的，对检索结果序列或检索结果页面本身的缓存都有助于大大减轻网络数据检索的负担。

面向网络信息的检索模型研究一直是网络信息相关研究的重点，研究的主要方法一般集中在以下两点：改进传统文本信息检索技术以适应网络信息需求；不对传统检索技术进行改

动，而尝试对传统检索结果与网络信息质量评估结果进行融合。前者主要的研究工作包括对网络半结构化文本检索方式的研究和对链接文本检索方式的研究等；而后一方面主要的研究工作则集中在将网络链接关系分析算法（如 PageRank）的结果融合到传统文本检索结果模型的尝试。而这方面最重要的研究平台之一是 TREC 的网络信息检索（Web Track）和超大规模文本检索（Terabyte Track）论坛。上述不少研究所基于的实验数据都来自于这两个论坛从 2001 年以来的积累，我们研究中涉及到的一部分实验结果也是基于这个研究领域比较公认的平台得出的。

总体而言，面向用户的查询服务技术是网络信息检索技术研究的热点之一，而主要的研究方向，则是基于传统文本信息检索技术，融入网络信息检索的应用需求元素，借以建立新的检索模型或对原有模型进行改进。当前，这方面研究集中在对用户信息需求的挖掘与发现、改进网络信息检索系统的用户交互方式等方面上。

3. 信息处理模块

除了上述两个模块之外的部分，都可以归入信息处理模块的范畴，将纷繁复杂的网页数据整理成可以被检索系统高效、可靠、方便使用的格式是这一模块的工作。具体来讲，信息处理包括对传统内容信息的处理和对非内容信息（主要是链接结构信息和文本结构信息等）进行处理两个方面的内容。

对文本内容信息的处理与传统文本信息检索系统的信息处理模块功能类似，最主要的目的是建立以词项（term）为中心的文本倒排索引，以便于在检索系统中使用检索功能模块进行内容相似度的计算，提高系统检索效率。这方面的主要研究工作集中在对倒排索引结构合理设计上，如是否记录文字位置信息以及记录的形式等。对中文信息检索系统而言，在这一部分还研究其特殊的预处理过程，即分词并建立词项的过程。由于这方面的研究有很多涉及到实际工程层面的实现原理与技术细节，因此相关研究工作的介绍并不多，仅有部分工程技术报告涉及，如 863 信息检索评测的相关工作等。

与纯文本数据相比，网络数据不仅包括非结构化的内容信息，也包括一定程度的结构化信息，这些结构化信息在网络信息检索工具评价数据质量、挖掘数据相关性等方面发挥着十分重要的作用。

因此，网络信息检索中对非内容信息的处理是其与传统文本信息检索系统最大的区别所在，而其中应用最广泛的是利用超链接结构分析方法的网络数据质量评价技术，因此在下面的介绍中将对这方面技术的主要发展情况进行讲解。

2.4.5　搜索引擎页面分析

对 SEO（搜索引擎优化）而言，让网站内的页面能够及时、全面地被搜索引擎索引、收录应该说是其的首要任务，这是实施其他 SEO 策略的最基本保证。不过，这也往往是容易被高估的一个环节，如我们时常可以看到某些人宣称自己的网站被 Google 收录了多少页面有几 KB 甚至几十 KB 等以证明 SEO 工作的成功。但客观地说，网页仅仅被搜索引擎索引、收录

是没有太大的实际意义，往往只能沦为浩如烟海的互联网世界中的殉葬品，更重要的是如何让网页出现在针对特定搜索项的 SERP（搜索结果页面）前几页。大部分人认为，让网站内尽可能多的页面被收录进搜索引擎索引数据库终归不是一件坏事，网页越多，曝光的机会也便越大，虽然最终效果存在疑问。

无论如何如果在对网站实施 SEO 时将重点放在网页被索引、收录的速度与效率，当然也无可厚非，而要想实现这一点，需要我们对搜索引擎如何收录、索引网页的机制有所了解。下面以 Google 为例，来介绍搜索引擎收录、索引网页的过程，希望能对读者能有所帮助。对其他搜索引擎如 Yahoo!、Live 搜索及百度而言，尽管在具体细节上存在差别，不过，基本策略都是类似的。

1. 收集待索引网页的 URL

在互联网上存在的网页数量绝对是个天文数字，每天新增的网页也不计其数，搜索引擎需要首先找到要索引收录的对象。

具体到 Google 而言，虽然对 GoogleBot 是否存在 DeepBot 与 FreshBot 的区别存在争议，至于是否叫这么两个名字更是众说纷纭，当然，名字本身并不重要至少到目前为止，主流的看法是，在 Google 的 robots 中，的确存在着部分专门为真正的索引收录页准备“素材”的 robots，这里我们姑且仍称之为 FreshBot。

它们的任务便是每天不停地扫描 Internet，以发现并维护一个庞大的 URL 列表供 DeepBot 使用，换而言之，当其访问、读取其一个网页时，目的并不在于索引这个网页，而是找出这个网页中的所有链接。

当然，这样似乎在效率上存在矛盾，有点不太可信。不过，我们可以简单地通过以下方式判断：FreshBot 在扫描网页时不具备“排它性”，也即是说，位于 Google 不同的数据中心的多个 robots 可能在某个很短的时间周期，如一天甚至是一个小时，访问同一个页面，而 DeepBot 在索引、缓存页面时则不会出现类似的情况，即 Google 会限制由某个数据中心的 robots 来完成这项工作的，而不会出现两个数据中心同时索引网页同一个版本的情况，如果这种说法没有破绽的话，则似乎可以从服务器访问日志中时常可以看到源自不同 IP 的 GoogleBot 在很短的时间内多次访问同一个网页以证明 FreshBot 的存在。因此，有时候发现 GoogleBot 频繁访问网站也不要高兴得太早，也许其根本不是在索引网页而只是在扫描 URL。

FreshBot 记录的信息包括网页的 URL、Time Stamp（网页创建或更新的时间戳），以及网页的 Head 信息（注：这一点存在争议，也有不少人相信 FreshBot 不会去读取目标网页信息的，而是将这部分工作交由 DeepBot 完成。不过，笔者倾向于前一种说法，因为在 FreshBot 向 DeepBot 提交的 URL 列表中，会将网站设置禁止索引、收录的页面排除在外，以提高效率，而网站进行此类设置时除使用 robots.txt 外还有相当部分是通过<mata>标签中的“noindex”实现的，不读取目标网页的 head 似乎是无法实现这一点的），如果网页不可访问，如网络中断或服务器故障，FreshBot 则会记下该 URL 并择机重试，但在该 URL 可访问之前，不会将其加入向 DeepBot 提交的 URL 列表。

总的来说，FreshBot 对服务器带宽、资源的占用比较少。最后，FreshBot 对记录信息按不同的优先级进行分类，向 DeepBot 提交，根据优先级不同，主要有以下几种。

A：新建网页。

B：旧网页/新的 Time Stamp，即存在更新的网页。

C：使用 301/302 重定向的网页。

D：复杂的动态 URL，如使用多个参数的动态 URL，Google 可能需要附加的工作才能正确分析其内容。（随着 Google 对动态网页支持能力的提高，这一分类可能已经取消）。

E：其他类型的文件，如指向 PDF、DOC 文件的链接，对这些文件的索引，也可能需要附加的工作。

F：旧网页/旧的 Time Stamp，即未更新的网页。注意，这里的时间戳不是以 Google 搜索结果中显示的日期为准，而是与 Google 索引数据库中的日期比对。

G：错误的 URL，即访问时返回 404 回应的页面。

优先级按 A~G 的顺序排列，依次降低。需要强调的是，这里所说的优先级是相对的，如同样是新建网页，根据指向其的链接质量、数量的不同，优先级也有着很大的区别，具有源自相关的权威网站链接的网页具有较高的优先级。此外，这里所指的优先级仅针对同一网站内部的页面，事实上，不同网站也有着不同的优先级，换而言之，对权威网站中的网页而言，即使其最低优先级的 404URL，也可能比许多其他网站优先级最高的新建网页更具优势。

2. 网页的索引与收录

接下来才进入真正的索引与收录网页过程。从上面的介绍可以看出，FreshBot 提交的 URL 列表是相当庞大的，根据语言、网站位置等不同，对特定网站的索引工作将分配至不同的数据中心完成。整个索引过程，由于庞大的数据量，可能需要几周甚至更长时间才能完成。

正如上文所言，DeepBot 会首先索引优先级较高的网站/网页，优先级越高，出现在 Google 索引数据库及至最终出现在 Google 搜索结果页面中的速度便越快。对新建网页而言，只要进入到这个阶段，即使整个索引过程没有完成，相应的网页便已具备出现在 Google 索引库中的可能，相信许多朋友在 Google 中使用 site:somedomain.com 搜索时常常看到标注为补充结果只显示网页 URL 或只显示网页标题与 URL 但没有描述的页面，此即是处于这一阶段网页的正常结果。当 Google 真正读取、分析、缓存了这个页面后，其便会从补充结果中逃出而显示正常的信息。当然，前提是该网页具有足够的链接，特别是来自权威网站的链接，并且，索引库中没有与该网页内容相同或近似的记录（Duplicate Content 过滤）。

对动态 URL 而言，虽然如今 Google 宣称在对其处理方面已不存在障碍，不过，可以观察到的事实仍然显示动态 URL 出现在补充结果中的几率远大于使用静态 URL 的网页，往往需要更多、更有价值的链接才能从补充结果中逸出。

而对于上文中所提及到的“F”类，即未更新的网页，DeepBot 会将其时间戳与 Google 索引数据库中的日期比对，确认尽管可能搜索结果中相应页面信息未来得及更新但只要索引

了最新版本即可考虑网页多次更新、修改的情况；至于“G”类即 404 URL，则会查找索引库中是否存在相应的记录，如果有，将其删除。

3. 数据中心间的同步

之前介绍过，DeepBot 索引某个网页时会由特定的数据中心完成，而不会出现多个数据中心同时读取该网页，分别获得网页最近版本的情况，这样，在索引过程完成后，便需要一个数据同步过程，将网页的最新版本在多个数据中心得到更新。

这就是之前著名的 Google Dance。不过，在 BigDaddy 更新后，数据中心间的同步不再集中在特定的时间段，而是以一种连续的、时效性更强的方式进行。尽管不同数据中心间仍存在一定的差异，但差异已经不大，而且，维持的时间也很短。

提高搜索引擎索引收录网页的效率，根据上面的介绍，可以看出，要想让你的网页尽可能快、尽可能多地被搜索引擎收录，至少应从如下几方面进行优化。

（1）提高网站反相链接的数量与质量，源自权威网站的链接可以让你的网站/网页在第一时间内被搜索引擎“看到”。

当然，这也是老生常谈了。从上面的介绍可以看出，要提高网页被搜索引擎收录的效率，首先要让搜索引擎找到你的网页，链接是搜索引擎找到网页的唯一途径，“唯一”一词存在些许争议，见下面的 SiteMaps 部分从这个角度看，向搜索引擎提交网站之类的作法没有必要且没有意义，相对而言，要想让你的网站被收录，获得外部网站的链接才是根本，同时，高质量的链接也是让网页步出补充结果的关键因素。

（2）网页设计要秉持“搜索引擎友好”的原则。

从搜索引擎 spider 的视角设计与优化网页出发，确保网站的内部链接对搜索引擎“可见”，相对于获得外部网站链接的难度，合理规划的内部链接是提高搜索索引与收录效率更经济、有效的途径，除非网站根本未被搜索引擎收录。

（3）如果你的网站使用动态 URL，或者导航菜单使用 JavaScript，则在网页收录方面遇到障碍时，应首先从这里着手。

（4）使用 SiteMaps。

事实上，许多人认为 Google 取消了 FreshBot 的一个主要原因便是 SiteMaps（xml）协议的广泛应用，认为这样只须读取网站提供的 SiteMaps 便可得到网页的更新信息，而不须 FreshBot 耗时费力地扫描。这种说法还是有一定道理的，虽然目前不能确定 Google 究竟直接使用 SiteMaps 作为 DeepBot 的索引列表还是用做 FreshBot 的扫描路标，但 SiteMaps 能够提高网站索引收录的效率则是不争的事实。如 SEO 探索曾做过如下的测试。

- 两个网页，获得的链接情况相同，一个加入 SiteMaps 而另一未加入，出现在 SiteMaps 中的网页很快被收录，而另一个页面则在过了很长时间后才被收录。
- 某个孤岛页面，没有任何指向其的链接，但将其加入 SiteMaps 一段时间后，同样被 Google 索引，只不过，出现在补充结果中。

当然，尽管网页未出现在 SiteMaps 中但仍能被 Google 索引也可以看出，Google 仍然使

用 FreshBot 或类似 FreshBot 的机制，当然这也容易理解，毕竟仍然有那么多未使用 SiteMaps 的网站，Google 不可能将其拒之门外。

4. 如何避免重复性收录

在互联网中，信息的重复性是在所难免的。然而，搜索引擎怎样识别重复信息？怎样判断哪些网页的信息是原创的？哪些是复制的？怎样认为哪些重复的信息是有价值的？哪些又是可以舍弃的呢？

在网站中，重复信息主要包括转载和镜像内容两大类。搜索引擎对页面分析的时候，必须具备识别重复信息的能力。因为大量的重复信息不但占用巨大的服务器资源，而且还增加了用户寻找信息的时间，降低了用户体验。但这并不意味着所有重复信息都是没有用的。搜索引擎会认为转载内容没有原创重要，赋予原创内容更高的权重，而镜像内容则几乎忽略。

转载页面时指与原创内容相同或相近的页面。然而，搜索引擎如何识别转载页面呢？首先，它把网页正文内容分为 n 个区域进行比较，如果其中有 m 个区域是相同或相似的，则认为这些页面是互为转载页面。

在确定页面的转载关系后，接下来，搜索引擎在结合页面的最后修改时间，页面权重等因素判断原创页面与转载页面。

镜像页面是指内容完全相同的两个页面。和上述一样，把内容分出 n 个区域，比对后 n 个区域完全一样则互为镜像页面。

镜像网站。狭义上是完全相同的网站，形成镜像网站主要有两种情况。一是多个域名或 IP 指向同一服务器的物理目录。另外就是整个网站内容被复制到使用不同域名或者不同 IP 的服务器上。

为了识别站点间是否互为镜像网站搜索引擎首先判断这些网站的首页是否互为镜像。然后综合页面权重，建立时间等诸多因素识别源网站。这也是为什么搜索引擎对于镜像网站收录极少，甚至不收录的原因。

2.5 知名的搜索引擎

为了提高搜索引擎的友好性，学习搜索引擎优化必须要熟悉各主要搜索引擎的习性。下面介绍一些知名的搜索引擎。

2.5.1 Google 搜索引擎

1996 年，斯坦福大学的研究生拉里・佩奇（Larry Page）和塞吉・布林（Sergey Brin）开发了一个名为 BackRub 的搜索引擎，该搜索引擎会使用链接确定各个网页的重要程度。到

1998 年，他们的工作正式启动并创办了一家公司，也就是现在的 Google（http://www.google.com），如图 2-9 所示



图 2-9　谷歌搜索引擎主页面

Google 也是最优秀的支持多语种的搜索引擎之一。提供网站、图像、新闻组等多种资源的查询。包括中文简体、繁体、英语等 35 个国家和地区的语言的资源。

2.5.2　百度中文搜索引擎

全球最大中文搜索引擎（http://www.baidu.com）。提供网页快照、网页预览/预览全部网页、相关搜索词、错别字纠正提示、新闻搜索、Flash 搜索、信息快递搜索、百度搜霸、搜索援助中心，如图 2-10 所示。



图 2-10　百度搜索引擎主页面

2.5.3　新浪搜索引擎

互联网上规模最大的中文搜索引擎之一（http://search.sina.com.cn）。设大类目录 18 个，子目录 1 万多个，收录网站 20 余万个。提供网站、中文网页、英文网页、新闻、汉英辞典、软件、沪深行情、游戏等多种资源的查询，如图 2-11 所示。



图 2-11 新浪搜索引擎主页面

2.5.4 雅虎中国搜索引擎

Yahoo!是世界上最著名的目录搜索引擎（http://cn.yahoo.com/）。雅虎中国于 1999 年 9 月正式开通，是雅虎在全球的第 20 个网站。Yahoo！目录是一个 Web 资源的导航指南，包括 14 个主题大类的内容， 如图 2-12 所示。



图 2-12 雅虎搜索引擎主页面

2.5.5 搜狗

搜狗搜索是搜狐公司于 2004 年 8 月 3 日推出的互动式中文搜索引擎（www.sogou.com/），成为全球首个中文网页收录量达到 100 亿的搜索引擎（目前已达到 500 亿以上），如图 2-13 所示。



图 2-13 搜狗搜索引擎主页面

2.5.6 搜搜

搜搜作为腾讯旗下的搜索引擎网站于 2006 年 3 月正式发布并开始运营，2009 年 9 月 3 日（http://www.soso.com/），搜搜搜索结果页面已经去掉“以下结果由 Google 提供”字样，

这很可能表明，搜搜已改用自主研发的搜索引擎技术。而之前，一直是由 Google 提供技术支持，如图 2-14 所示。



图 2-14　搜搜搜索引擎主页面

2.5.7　网易有道搜索引擎

网易自主研发的全新中文搜索引擎（http://www.youdao.com/），2006 年底有道推出测试版，并于 2007 年 12 月 11 日推出正式版。目前有道搜索已推出的产品包括网页搜索、图片搜索、音乐搜索、视频搜索、博客搜索、购物搜索、网易返现、有道购物助手、有道词典、有道手机词典、有道翻译、网易八方、有道笔记、工具栏、网址导航、有道阅读、有道热闻等，如图 2-15 所示。



图 2-15　有道搜索引擎主页面

2.6　小结

搜索引擎优化的内容包括很多方面，我们做搜索引擎优化的宗旨就是让搜索引擎喜欢。当然，搜索引擎程序的排名规则也会变化，这需要不断地观察和研究，可能从容地适应各种变化。作为一个 SEO 工作者，必须深入理解搜索引擎的工作原理，只有这样才将能 SEO 工作做得更完善，以达到提高网站的搜索引擎友好性。

第 3 章　关键词

想要对网站进行搜索引擎优化，肯定先要为网站的页面选择适当的关键词，之后围绕着核心关键词展开相关的工作以让关键词能在搜索引擎排名中达到满意的结果。而很多人在选择关键词的时候，往往选择热度非常高，搜索量非常大的词来做关键词，吸引流量。这个方法无可厚非，但是，使用这种方法也是要讲究方法和原则的。我们选择热度关键词的时候，不要脱离了自身网站的主题内容。本章将给大家具体介绍如何正确选择关键词的方法。

3.1　什么是关键词

我们常说的关键词其实就是你输入搜索框中的文字，也就是你命令搜索引擎寻找的东西。你可以命令搜索引擎寻找任何内容，所以关键词的内容可以是宠物、人名、网站、新闻、小说、软件、游戏、星座、工作、购物、论文等，如图 3-1 所示。



图 3-1　搜索关键词

首先，关键词可以是中文、英文、数字，或中文英文数字的混合体。

例 1：“金毛寻回犬”、“windows”、“119”、“F-1 赛车”。

其次，关键词可以输入一个，也可以输入两个、三个、四个，甚至可以输入一句话。

例 2：你可以搜索“爱”、“美女”、“mp3 下载”、“游戏 攻略 大全”、“蓦然回首，那人却在灯火阑珊处”。

输入多个关键词搜索，可以获得更精确、更丰富的搜索结果。

例如，在百度搜索引擎中搜索“牧羊犬”，可以找到几万篇资料；而搜索“比利时牧羊

犬”，则只有严格含有“比利时牧羊犬”连续 6 个字的网页才能被找出来，不但找到的资料只有几百篇，而且资料的准确性也比前者高很多。

因此，当你要查的关键词较为冗长时，建议将它拆为几个关键词来搜索，词与词之间用空格隔开，如图 3-2 所示。



图 3-2　关键词搜索模式

3.2 关键词的选择

关于网站核心的关键词选择不能马虎，必须经过细致地审查才能保证这个关键词确实有网民在搜索，没人搜索的关键词是没有任何价值的。

对 SEO 没有概念的人决定目标关键词时首先想到公司名称或自己的产品名称。但是当企业或网站没有品牌知名度时，没有用户会搜索公司名或网站名。产品名称如果不包含产品的通用名称，也往往没人搜索。

有很多时候即使是使用此行业中最通用的词语，也不一定有足够的真实搜索数量，最典型的就是 SEO 这个词本身。通过百度指数我们看到 SEO 每天的搜索量有八九千次以上，首先这与很多其他真正用户在搜索的词相比，已经属于比较低的。其次，搜索 SEO 的人很多是站长在研究记录 SEO 这个词排名的新动向，而不是对 SEO 服务感兴趣。

要确定适当的关键词，首先要做的是，确认用户搜索次数达到一定数量级。在这方面做出错误的方向选择，对网站的影响将会是灾难性的。经过核心关键词确定与关键词扩展，应该已经得到一个至少包含几百个相关关键词的大列表。这些关键词需要合理分布在整个网站上。

在确定了核心关键词后，接下来就是进行关键词扩展。对一个稍有规模的网站来说，几个关键词远远不够，还需要找出比核心关键词搜索次数少一些的跟多关键词，安排到分类或频道首页，从而扩展出几百甚至几千个关键词都很常见。

3.2.1 降低优化难度

即使我们找到了有一定搜索量的关键词，那也不能说明我们就一定要把目标选择在这些最热门、搜索次数最多的关键词上。虽然搜索“新闻”、“律师”、“租房”、“机票”、

“减肥”、“旅游”、“化妆品”等这些词的用户很多，但是对中小企业和个人站长来说，要把这些词做到前几位，难度是非常高的。可以说没有强大的资源和人力支持，是不可能做到的。做关键词研究就是要找到被搜索次数比较多，同时难度不太大的关键词，网站优化才有可能在一定的预算、周期下取得较好的效果。

3.2.2 寻找有效流量

排名和流量都不是目的，流量带来的转化才是目的。就算公司有足够的实力将一些非常热门的关键词排到前面，也不一定是投入产出比最好的选择。

假设网站发布的内容信息是“律师服务”，然后将网站核心关键词选为“律师”，这样的情况一般来说并不会是最好的选择，因为用户搜索“律师”这个词的动机和目的很难判定。用户有可能是在寻找律师服务，但也可能是在寻找律师资格考试内容，也可能是在找大学专业报考指导，这样的用户来到提供律师服务的网站就没有什么机会转化为客户。

如果把核心关键词定为“北京律师”针对性就强得多，用户已经透露出一定的意向；如果目标关键词定为“北京刑事律师”，则用户意向就非常明确，几乎可以肯定这个用户是在寻找特定的律师服务，这样的搜索用户来到你的网站，转化为客户的可能性将大大提高。

3.2.3 内容相关

目标关键词必须与网站内容有关联。SEO 早期曾经在页面上堆积搜索次数，但都是与本网站没有实际相关的关键词，也曾经起到很好的作用，能带来不少流量。不过，现在这样的做法早已过时，网站需要的不仅仅是流量，更是有效流量，靠欺骗性的关键词带来访客不能完全转化，对网站毫无意义。这样的排名和流量不是资产，而是负担，除了消耗带宽，没有其他作用。

假设如果你的网站想要卖“化妆品”或者“减肥产品”，就不要想着靠“汽车”或“房子”这种关键词带来流量。抛开关键词的优化难度和可能性不谈，就算搜索这种不相关关键词的访客来到网站也不会买你的产品或服务。

当然，这不一定适用于某些网站，如新闻门户或纯粹依靠广告营利的信息类网站。很多门户类网站包罗万象，内容相关性判断也比较模糊。对这些网站来说，只要有流量，就有一定的价值。

3.2.4 搜索次数多、竞争小

在做关键词优化时，应寻找的最好的关键词是那些搜索次数最多、而且竞争程度最小的词，这样既保证 SEO 代价最低，又保证流量最大。可惜现实不是这么理想的，大部分搜索次数多的关键词，也是竞争大的关键词。不过，通过大量细致的关键词挖掘、扩展，列出搜索次数及竞争程度数据，还是可以找到搜索次数相对多、竞争相对小的关键词。

研究搜索次数比较直接、简单，Google 关键词工具及百度指数都提供搜索次数数据，详情请参考 SEO 工具相应章节。

要确定一个词的竞争强度是比较复杂的，这需要参考查询的数据很多，而且很多带有不确定性。根据搜索次数和竞争程度可以大致判断出关键词效能。在相同投入的情况下，关键词效能高的词获得排名的可能性较高，可以带来更多流量。

例如：

1. 主关键词不可太宽泛

这实际上是上面两点的自然推论。关键词宽泛，竞争太大，所花代价太高，搜索词意图不明确，转化率也将降低。如做房地产的公司，想当然地把“房地产”作为目标关键词；做旅游的公司就把“旅游”作为目标关键词，这都犯了主关键词过于宽泛的错误。

2. 主关键词不可太特殊

选择主关键词也不能走向另外一个极端。太特殊、太长的词，搜索次数将大大降低，甚至没有人搜索，不能作为网站主关键词。

如果说“律师”这个词太宽泛，那么选择“北京律师”比较适当，根据不同公司业务范围，可能“北京刑事律师”更合适。但是如果选择“北京新街口律师”就不靠谱了。这种已经不属于长尾关键词可以考虑以内页优化。

特殊的关键词还包括公司名称、品牌名称、产品名称等。

综上所述，网站主关键词、或者称为网站核心关键词，既不能太长、太宽泛，也不能太短、太特殊，需要找到一个平衡点，才能达到想要的效果。

3.2.5 商业价值

选择具有商业价值的关键词，因为不同的关键词其具有的商业价值也不同，而且就算长度等同，也同样会导致不同的转化率。

如搜索“发电机原理”的用户购买意图就比较低，商业价值也低，他们很可能是在做研究，学习发电机知识而已。而搜索“发电机图片”的用户商业价值有所提高，很可能是在寻找、购买发电机的过程中想看看产品实物是什么样。搜索“发电机价格”，购买意图大大提高，已经进入产品比较选择阶段。而搜索“发电机促销”或“发电机购买”，其商业价值更是进一步提高，一个大减价信息就可能促成用户做出最后的购买决定。

站长在做关键词研究的时候，是可以通过各种方式手段查询到大量搜索词的，通过常识就应该能够判断出不同的词的购买可能性。购买意图比较强烈、商业价值较高的关键词应该是优化时最先考虑的，无论内容规划，还是内部链接安排，都要予以侧重。

搜索结果中页面右侧、最上和最下面有多少个广告的显示结果，也是衡量竞争程度的指标之一。

一般来说广告商内部有专业人员做关键词研究和广告投放，他们必然已经做了详细的竞

争程度分析及营利分析，只有能产生效果和营利的关键词，才会去投放广告。如果说搜索结果数还只是网上内容数量带来的竞争，竞价数则是拿着真金白银与你竞争的真实存在的竞争对手数目。

搜索结果页面的右侧最多可显示 8 个广告。比较有商业价值的关键词，通常都会显示满 8 个广告结果。如果某个关键词搜索页面右侧只有两三个广告，说明关注这个词的网站还比较少，竞争比较低。

要注意的是，竞价数需要在白天工作时间查看，广告商投放竞价广告时，经常会晚上停止广告。对于经常晚上工作的 SEO 人员来说，如果半夜查看搜索结果页面，没有看到几个广告商，就认为没有多少人参与竞价，这样很可能导致误判。

3.2.6 竞价价格

几大搜索引擎都有提供相关查询工具，可以让广告商投放前就能清晰明了的看到某个关键词的大致价格，这样的价格可以排到第几位，以及能带来多少点击流量。如 Google 的关键词流量估算工具。

显然，竞价价格越高，竞争程度也越高。当然也不能排除两三个广告商为了争抢广告位居第一名而掀起价格战，把本来竞争程度不太高的关键词推到很高的价格。实际上如果广告商只想出现在第四、第五位的话，竞价价格大幅下降，参与竞价的广告商人数也没有那么多。

另外，某些利润率高的行业，搜索竞价经常超出自然排名的真正竞争程度，如律师服务、特效药品。这些行业产品及服务的特性决定了一个订单的利润可能是成千上万，企业可以把竞价提到相当高的程度，甚至一个点击几十元钱也不罕见。而销售书籍、服装、化妆品等，利润不是很高，竞价价格也不可能太高。广告价格的巨大差异，并不能说明自然搜索竞争程度真的有这么大差别。

3.2.7 各种形式的变体

用户都是搜索什么样的关键词来到我们的网站的呢？我们在查看网站现有的流量时，经常可以看到一些 SEOER 自己并没有想到的关键词流量出现。用户之所以能搜索这些关键词找到网站，说明搜索引擎认为你的网站与这种关键词有比较高的相关性。把这些关键词输入到 Google 工具，生成更多相关词，也是一个很好的关键词扩展方式。

1. 同义词

假设核心关键词是酒店，与酒店基本同义的还有饭店、旅馆、住宿、旅店、宾馆等。再如网站推广、网络推广、网络营销意义也很相近。

2. 相关词

虽然不同义，作用却非常类似的词。如网站建设、网页设计、网络营销与 SEO 非常相关，目标客户群也大致相同。

3. 简写

如 Google PR 与 Google PageRank，北京大学与北大。

4. 错字

还有一类变体是错字。如艾滋病与爱滋病、SEO 每天一贴与 SEO 每天一帖、点石与电石。有不少用户使用拼音输入法经常会输入错字、同音字，所以有一些搜索量。但是优化错别字就不可避免地要在页面中出现这些错别字，可能会给网站用户带来负面观感，使用时需要非常小心。

3.2.8 补充说明文字

如果前面已经找到了几百个的关键词，那么把这些词交叉组合起来又很容易生成数千个扩展关键词。这些比较长的组合起来的关键词看起来搜索次数并不多，但数量庞大，累计起来能带来的流量潜力也非常可观。

1. 地名

有的核心关键词配合地名很明显，如：

- 旅游：云南旅游、海南旅游、北京旅游。
- 酒店：上海酒店、南京酒店。

有的关键词看似与地理位置无关，却有不少用户会加上地名搜索，如：

- 北京办公家具、上海办公家具、广州办公家具。
- 上海鲜花、深圳鲜花。

哪种核心关键词配合地名有搜索量，需要查看搜索引擎关键词工具。有的关键词就完全没有人加上地名搜索，如减肥。

2. 品牌

核心关键词加上品牌名称很常见，如：

- 电视机：康佳电视机、索尼电视机。
- 手机：诺基亚手机、摩托罗拉手机、索爱手机。

3. 限定和形容词

例如：

- 主机：免费主机、免费服务器。
- 电视机：电视机价格、电视机促销。
- 京东商城：京东商城官方网站。

以上提到的免费、促销、价格、官方网站这些附加限定词都很常见。便宜、怎样、是什么、好吗、评测、图片等也都是常见限定词。

4. 单词交叉组合

上述提到的核心关键词、同义词、近义词、相关词、简写、地名品牌限定词等，放在一起又可以交叉组合出很多变化形式。如北京办公家具价格、京东商城电视机促销、诺基亚手机评价、云南旅游攻略等。

关键词的选择是 SEO 最具技巧性的环节之一。只有选择正确的关键词，才能使网站 SEO 走在正确的大方向上，关键词的选择决定了网站内容规划、链接结构、外部链接建设等重要的后续工作。

3.3 关键词工具

关键词查询工具是指可以提供指定关键词的搜索量及相关信息的工具，在对网站进行优化时，适当的使用一些辅助工具不但可以节省大量时间，还可以帮助我们进一步提高网站的搜索引擎友好性。

3.3.1 站长工具箱

站长工具箱可以检索某个关键词的扩展关键词，百度指数，谷歌搜索频率，以及某个词在某个网站的排名状况，如我的 123 宠物网的关键词“宠物网”一词在百度、谷歌等搜索引擎检索的排名位置，非常实用，对独立网站使用比较适合，可以用它来扩展网站的相关关键词、长尾关键词等，如图 3-3 所示。



序号	关键字	指数	排名第一位
1	宠物狗	1483	www.58.com
2	宠物狗的种类	816	dog.aigou.com
3	宠物狗图片	437	dog.aigou.com
4	宠物狗的名字	339	www.ganji.com
5	宠物狗价格	317	www.58.com
6	宠物狗大全	267	dog.aigou.com
7	宠物狗品种	164	dog.aigou.com
8	宠物狗名字大全	163	www.ganji.com
9	宠物狗狗	155	pet.pclady.com.cn
10	宠物狗58同城网	123	www.58.com
11	宠物狗品种大全	117	www.petyoo.com
12	宠物狗种类	107	wenku.baidu.com
13	石家庄宠物狗	100	shijiazhuang.baixing.com
14	宠物狗网站	95	www.cndog.net
15	宠物狗大赛	89	v.youku.com

图 3-3　站长工具关键词挖掘

3.3.2　百度指数工具

百度指数工具功能齐全，检测长尾关键词、扩展词，检索速度快，可看出每日的搜索变化情况，还可提供相关资讯及相关的新闻动态，如图 3-4 ~ 图 3-6 所示。

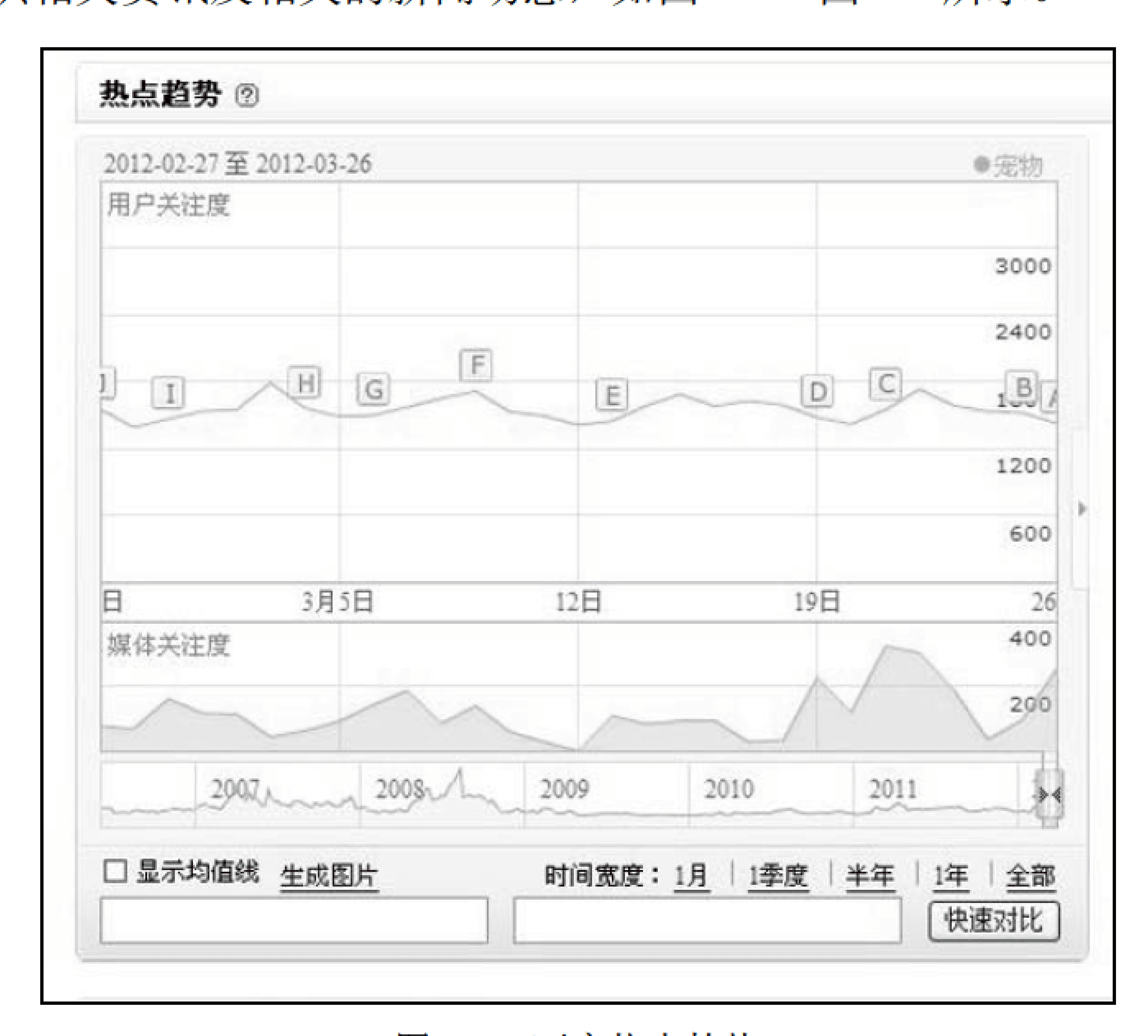


图 3-4　百度热点趋势



图 3-5　百度相关新闻



图 3-6　百度相关检索

3.3.3　搜索建议下拉框

搜索建议的下拉框是在百度或谷歌搜索栏中输入核心关键词时，搜索栏会自动弹出与此相关的建议关键词，如图 3-7 所示。



图 3-7　百度相关搜索下拉框

3.4　关键词密度

关键词密度（Keyword Density）也叫关键词频率（Keyword Frequency），所阐述的实质上是同一个概念，它是用来量度关键词在网页上出现的总次数与其他文字的比例，一般用百分比表示。相对于页面总字数而言，关键词出现的频率越高，那么关键词密度也就越大。举个例子，如果某个网页共有 100 个词，而关键词在其中出现 5 次，则可以说关键词密度为 5%。

当然，这个例子只是一种理想的简化方式，它没有有效包括 HTML 代码里面的诸如 <meta>标签中的 Title、Keywords、Description，图像元素的 Alt 文本、注释文本等，这些在计算关键词密度时也都要考虑在内的。同样地，在计算关键词密度时，也不仅仅指网页可见部分中关键词出现的次数，也包括在上文中提到的非可见部分中的出现次数。同时，也要考虑 Stop Words（停用词），这些词往往会在很大程度上稀释关键词密度。搜索引擎在算法上要比这复杂得多，但基本策略与此近似。

例如：

1. 中文关键词

在中文里，字或者词之间并不存在自然分隔符，而且中文里的词通常由两个或两个以上的中文字符组成。因此，搜索引擎不能借助分隔符对页面正文内容进行分词，而是按照某种算法把页面正文内容划分为若干个中文词汇。

2. 中文关键字密度

为了让读者可以更加容易理解中文关键字密度，下面举一个简单的例子进行说明。如网页内容为“我的笔记本”，搜索引擎将其切分为“我”、“的”、“笔记本”，则关键字“笔记本”在这个文档中的密度就是“1/3”。

对于短语关键字，通常通过计算组成短语里的每个词的密度去衡量该短语的密度。例如“智能手机走进大众市场”可以切分为“智能”、“手机”、“走进”、“大众”、“市场”。其中的“智能”及“手机”各出现了一次，则它们的关键字密度都是“1/5”。

3. 英文关键词密度

在英文中，同一单词的不同形式会被认为是两个不同的单词（大小写除外）。如 battery 与其复数形式 batteries，会被认为是两个不同的单词。

在英文书面表达里，空格是单词间的自然分隔符，而句点就是一个句子或者段落结束的标记。根据这个特征，搜索引擎就可以轻易地对网页正文内容进行准确的划分。

而对于词组，通常通过计算组成词组的单词的密度去衡量该词组的密度是否合理。

在每个页面确定了自己的关键词之后，需要在页面内容中适当出现这些关键词。关键词密度（Keyword Density）是用来量度关键词在网页上出现的总次数与其他文字的比例，一般用百分比表示。相对于页面总字数而言，关键词出现的频率越高，关键词密度也就越大。关键词密度一般在 2%~7%较为合适。不要进行关键词堆砌，即一页中关键词的出现不是根据内容的需要而安排的，而是为了讨好搜索引擎人为堆积关键词（Stuffing）。这已经被搜索引擎归入恶意行为（Spamming），有遭到惩罚的危险。

4. 英文 SEO 的关键词布局

英文 SEO 关键词布局的基本原则是：无所不在，有所侧重。

页面的正文内容当中必须适当出现关键词，并且“有所侧重”。从上到下——范指用户阅读习惯形成的阅读优先位置；从左至右——成为关键词重点分布位置，包括页面靠顶部、左侧、标题、正文前 200 字以内，在这些地方出现关键词对排名更有帮助。

要做英文 SEO 优化的网站进行域名选择和网页、文件夹命名时，也可以考虑包含关键字，对排名有一定作用。

上述这些就是英文 SEO 的关键词密度及布局，和中文 SEO 的关键词密度和布局没太大区别，所以，知道如何做中文 SEO 的关键词密度及布局，英文 SEO 也是手到擒来的事。

3.4.1 关键词密度控制

一直以来包括百度、谷歌、雅虎等这些搜索引擎都将关键字密度作为其排名算法考虑因素之一，每个搜索引擎都有一套关于关键字密度的不同的数学公式。合理的关键字密度可使你获得较高的排名位置，密度过大，反而会起到相反的效果。

1. 首页关键词密度

首先能放关键词的肯定是标题、关键词和描述，其次为显示替换类的，再就是文章内容。更新的文章、图片具体增加关键词的密度有：1 显示、2 替换、3 注释、4 文件名称、5 子目录名称。

首页密度控制在5%~10%之间，但具体控制在多少需要进一步进行分析。

（1）域名中是否带关键词简写拼音，如果带则可将密度调协得高点。

（2）域名是com的可以多一些，cn的则可少一些。

（3）是老域名还是新域名，刚开始的新域名最好控制在6%比较稳妥，老域名可以适当高点。

（4）关键词的热度，如果关键词很热的话，关键词密度最好低点，当然太低的竞争又上不去，可以先低，等排名上去了再增加关键词密度。

2. 栏目关键词密度

栏目关键词的方法基本和首页一样，它与增加关键词密度都是网页技术，不过要注意的是，首页和栏目页所采用的增加关键词密度的方法要有些不相同，这样才不会留下明显的痕迹。栏目页关键词应控制在6%以内。

基于上面的说法，控制栏目关键词密度需要注意的地方如下。

（1）栏目应该是真正的栏目，如栏目是做的“鲜花”的，那么内页也应该是有关鲜花的内容，内页关键词密度控制也应该是鲜花。

（2）栏目既然只是一个栏目，那么该栏目首先要注意的则是栏目方面的关键词，对首页的关键词都可以放置在后面。

（3）栏目关键词应该是和首页关键词相关。

（4）栏目关键词与首页关键词又有不同，因为是用栏目关键词来另外开辟一个排名，如做的是结婚类的网站，如果另外做个栏目叫做鲜花，这样就能在鲜花方面排名，排名显示的肯定是栏目不是顶级域名。

（5）如果网站不建议栏目做排名的话，那么得将所有页面都看作是子页面。

3. 子页面关键词密度

理论上子页面关键词密度是可以控制在10%而不被处罚的，但是如果一个网站所有页面都是10%的关键词密度的话，岂不是全站到处都是关键词了，这样的网站会被优化到别人看都不想看的状态。

控制子页面关键词密度要注意以下几点。

（1）必须是平均分配，不要有些6%，而有些只有1%，这样的话，一旦控制不好容易让有些页面超出密度要求。

（2）不能只是网站顶部和顶部来增加密度，必须有30%以上的关键词密度来自于页面的文章内容。

（3）子页面是居于网站之下，栏目之下，所以一定得跟着网站和栏目来设计关键词。

4. 关键词密度应该出现些错误化

（1）如做“汽车美容”的网站，栏目是“汽车美容的价格”，子页面的关键词要做的肯定是上面两个关键词，根据百度分词计算方式可以分为“汽车”、“美容”、“价格”、

“美容价格”、“汽车价格”、“汽车美容”、“车美容”、“车价格”等，具体分词方式要根据具体词来分，那么子页面可以对上述分词中的关键词分别进行密度增加，而不是单一的增加一个词，这样优化痕迹更不明显。

（2）经常也能出现些错误的关键词如“起车”、“加格”等，因为拼音是一样的，百度能识别，但是错误不会影响你的密度。

（3）当然如果能做 10~20 个页面，则没有关键词也可以。

3.4.2 关键字密度对网站排名的影响

就实施惩罚所容许的关键字密度的阈值而言，不同的搜索引擎之间也存在不同的容许级别。对过度优化如关键词 Spam 而言，不同的搜索引擎容忍的阈值也不尽相同，从大到小排列大致依次是 Google、MSN，然后是百度、Sogou、，IASK、Yahoo！。

1. 关键词密度的阈值

什么样的关键词密度为妙呢？关键词密度是一个模糊的概念，没有一个准确的公式来限定其密度。各家搜索引擎的密度值控制都不一样，甚至是同一家的搜索引擎，对不同网站的关键词密度的大小所能允许的容忍阈值也不相同。如同样一个页面，新浪、CCTV 的网站密度值达到 20%可能就没有什么事情，但如果是你的网站估计马上就被屏蔽了。这还涉及到了搜索引擎的信任值问题，在此先不作介绍。

关键词密度并不是越高越好。一般来说，在大多数的搜索引擎中，ing 关键词密度在 2%~8%是一个较为适当的范围，有利于网站在搜索引擎中排名，同时也不易被搜索引擎视为关键词 Spam。笔者认为在百度上的关键字密度对排名较为重要，5%左右的密度应该是比较理想的，但是也根据网站页面的字符数而定，Google 的可以更高些。很多相关工具对此都给出了大概的范围，如图 3-8 所示。

关键词	出现频率	2%≦密度≦8%
软件著作权	7	5.0%
著作权登记	5	3.6%
计算机软件著作权	5	5.7%
软件著作权登记	5	5.0%

图 3-8 站长工具关键词密度查询

2. 怎样查询网页关键词密度

在推荐上的“页面关键词密度查询”中可查询网页关键词的密度。

相对百度，Google 搜索引擎赋予关键词密度的权重更小了，很多排名靠前的网页关键词密度可能高到 20%以上，也可能完全没有关键词。有人曾经专门就这个观点做了一个实践，通过一个布满关键词的页面进行观察，百度更重视的是关键词密度，而 Google 更为重视的则

是外部链接的建立。

3. 什么是增加关键词密度的最好方式

毫无疑问，最佳选择的是通过网站本身的内容来做为有效支撑。所以只要按逻辑，按语法正常写你的网页，就不必太顾及关键词的密度。一个页面中出现关键词密度很高的情况，只要是实际需要的，便可以保留，毕竟是给你网站的访问者看的，而不是做给搜索引擎看的，尽量提供最有价值的信息给浏览者才是最重要的，不过需要澄清的是，那样的页面有可能会被搜索引擎自动的过滤掉，因为关键字过多很可能会触发关键字堆砌过滤器（keyword stuffing filter）。

3.4.3 如何布局关键词

要达到理想的网站优化效果，不仅仅要为我们的网站或网页选定恰当的、有效的关键词，更重要的是如何在网页中恰当地将这些关键词嵌入到内容中。

1. 关键词的位置

关键词出现在页面文件的哪个部分及先后顺序，对网页出现在 SERP 页面中排名的影响逐步降低，如下实例：

```
<!--页面注解-->
<html>
<head>
<title> 页面标题</title>
<meta http-equiv=Content-Language content=zh-cn>
<meta http-equiv="Content-Type" content="text/html; charset=gb2312">
<meta name="keywords" content="关键词">
<meta name="description" content="网站描述">
<link href="inc/css.css" _fcksavedurl=""inc/css.css""
_fcksavedurl=""inc/css.css""
type="text/css" rel="stylesheet">
</head>
<body>
<div>
<h1>页面内容标题</h1>
<h2>页面相关性标题</h2>
<h3>标题系列</h3>
<h4>标题系列</h4>
<h5>标题系列</h5>
<h6>标题系列</h6>
<img src="xxx.jpg" alt="图片说明">
<a href="/" title="链接说明">链接词</a>
<strong>重点关键词强调</strong>
```

```
<b>关键词强调</b><u>关键词强调</u><i>关键词强调</i>
</div>
<div>
版权部分关键词强调
</div>
</body>
</html>
```

注：在<body>…</body>标签区中的罗列，都是基础的 HTML 标签，这些标签出现的次序是可以变化的。

第 1 行<!--页面注解-->。这个标签是用来做 HTML 代码注解的，很多站在这里会写上关键词，这种形式的注解在浏览器中是不会显示的，而搜索引擎可以读到。所以这个标签有的时候会成为作弊的一种形式，不要过多的用这种标签来注明你的关键词，否则如果搜索引擎认为你是作弊的话，就惨了。

第 4 行<title>页面标题</title>。对于做百度 SEO 的朋友，一定要记得这个标签可能是你是否真正能够做到排名靠前的重点，而且这个重点绝对不容忽视。标题主要注意你的用词，如果你做垃圾站，堆关键词也无所谓；如果你要做正规站点，请你认真对待，写得标准一点，不要堆关键字，保持你的主要词在这里出现 3 次，太多对你没有好处，对排名也不会有好处。Google 并不喜欢太长的标题，百度也偶尔提到只希望在标题的前 20 个字里出现你的关键词。明确这些以后，还要注意的是，你的主要关键词应该在标题里越靠前越好。

第 7 行<meta name="keywords" content="关键词">。它的权重越来越低，但是再低都是有权重的，只要认真对待，这里的词是有效果的。这里需要注意以下两点：

- 大部分人在此处写的关键词不会超过 3 个，有的人直接只写一个，这是有好处的，每一处标签都有权重，你强调的内容越少，那么它的权重就会越高，所以不要写太多，否则不如不写，而让搜索引擎自己去判断你的页面关键词。
- 几乎所有的 SEO 专业人员都会跟你说尽量在每一个页面写不同的关键词，避免被认为是重复页面的危险。

第 8 行<meta name="description" content="网站描述">。这个主要是为了在搜索结果里出现索引结果，索引结果就是在百度搜索你的站的时候标题下面的摘要内容。这是一个页面的导读，用来说明网页的主要内容，当然现在的权重不高，跟 keywords 标签一样，不太受重视。而且现在由于搜索引擎机器人的越来越智能化，基本上已经实现整个页面的摘要内容都能随意出现在它的索引结果里，搜索引擎会根据你搜索的内容不同而显示不同的索引。你如果要写 description，就认真去写，很多人是用这里来控制搜索引擎的索引显示的内容，因为网站描述写得越具体，就越有可能在索引的时候出现这个标签中的内容。

第 13 行<h1>页面内容标题</h1>。这个标签是页面中最重要的标签，也是搜索引擎定位这个页面的主要依据，如果你的<h1>标签里没有出现你的页面关键字，那真是一种很大的损失，这里的权重相对于其他的在<body>区的标签来说更加重要。在 HTML 代码中，它的重要

性仅次于<title> 。很多人放弃写这个标签，这对于搜索引擎是很不友好的。这里可以不出现其他强调关键词的标签，但是这个标签一定要出现，这是告诉搜索引擎你的页面用一句话应该如何表达，如果这一行一个字都没有，再智能化的机器人可能都会以你的<title> 来决定是否用其作为一个词的排名，因为他们程序的原因，很多时候一个网站的页面标题都是完全一样的。如果你没有办法修改 title，已经损失很大了，而如果你的页面中的<h1>标签也没有出现的话，那么搜索引擎要怎样去理解这个页面呢？这对于 SEO 来说，是一件很糟糕的事情。所以，认真对待页面中的<h1>标签。需要注意的是，这个标签在同一个页面不要出现太多，道理很简单，太多会让别人无法确定你页面中哪一个才是最重要的，无法确定重点跟没有这个标签我觉得没有太大的分别。所以，最好只出现 1 次，以达到明确说明的 目的。

第 14 行<h2>页面相关性标题</h2>。这个标签与<h1>是一样的，但是权重相对来说较低，如果你的页面没有<h1>，那么这个标签对于页面来说基本上就相当于<h1>的重要性了。这里面你包含的是长尾关键词相关的内容，如你的页面做几个关键词，主词应该出现在<h1>里，那么其他的词以及你想发展的词就可以出现在<h2>里，这样来说是一种合理的安排。而<h3>~<h6>标签，其重要性是依次递减的，在页面中的权重越来越低。

第 19 行<img src="xxx.jpg" alt="图片说明">。我想这个标签很多人都知道，它的作用很明确，不是指定一张图片的说明文字，让搜索引擎知道这张图片是关于什么的。如你的图片是"刘亦菲"，那么你的 alt="图片说明"中的"图片说明"就应该是与"刘亦菲"相关的词或者句子。这样写当然也有利于你在图片搜索的结果里显示排名更前面一点，不过现在百度基本上是以页面标题作为一张图片排名的要素，当然，你的 alt 标记做得好，将会增加你的图片匹配度。

第 20 行链接词。这是链接说明，同时也是一个页面链接的 SEO 重点，这里的“链接词”相当的重要，如果你跟别人交换链接，你这个词的写法是起着外链好坏与否的决定性因素的。“链接词”建议最好写成自己的关键词，如果觉得实在不行，则一定要在“链接说明”里出现你的关键词。否则对于排名，你的链接做得没有太大的好处。这在 SEO 专业术语里还有一种叫法，叫做“Google 炸弹”，就是用很多站点，同时链接到一个网站，然后所有的“链接词”都写一个词“刘德华”，不久以后 Google 就会将这个站当做“刘德华”这个词最权威的站点，即使被那些站点链接的这个页面没有任何与“刘德华”相关的内容，这个网站在“刘德华”这个词的排名情况也会很好。不过 google 已经调整了算法，但是这样做链接肯定是有用的，一定要认真对待你的链接。还有一个重点，就是对方的链接，如果是 JS 调用的，你就别想在他那里得到任何权重，因为搜索引擎不会去读取 JS 里的内容，所以，建议大家好好做链接，不要你的站做了链接却跟没做一样。像百度空间的友情链接就是 JS 文件调用的，所以百度空间里的友情链接你是没有权重的。具体可以查看 HTML 代码，正常情况下，在一个页面中的<a>...标签内找不到你的网站地址，基本上这就是属于无效链接。

第 21 行重点关键词强调。这个标签的重要性仅次于<h1>，它比<b>、<u>、<i>等标签要权重高。所以如果你的页面有很重要的词，而你想强调它，则首先选这个

<strong>标签，接下来相对不重要的词则可选<b>、<u>、<i>等标签。当然，不要用得太多了，用得太多没有重点就等于没有用了。

第 26 行版权部分关键词强调。这个权重越来越高了，在底部版权部分，很接近版权部分出现的关键词，是一种 SEO 策划的手法。在这里强调关键词，是让关键词布局更加合理，而不是头部出现关键词，中部强调很多次，而底部却什么也没有关联。

2. 关键词密度或关键词频率

在网页中关键词出现的频率越高，搜索引擎便会认为该网页内容与相应关键词的相关性更高，从而更易出现在 SERP 页面的前端。

3. 关键词 URL 对应表

关键词分组完成后，建议 SEO 工作人员将关键词搜索次数，目标 URL 等情况列表，如表 3-1 所示。

表 3-1 关键词 URL 统计列表

关键词	百度指数	目标 ULR	是否收录	目前排名	目前月搜索流量
宠物网	757	http://www.123pet.com.cn/	是	3	2615
宠物论坛	232	http://bbs.123pet.com.cn/	是	2	1423
马犬	855	http://www.123pet.com.cn/gou/maquan	是	3	3456

3.5 预估搜索流量的价值

流量并不是目地，促成交易获得利润才是目标。获得预计流量后，结合网站转化率及平均订单销售额和平均每单利，就可以计算出通过 SEO 获得的搜索流量能给公司带来多少实际价值。

如预计搜索流量为每月 8 万个独立 IP，网站历史平均转化率为 1%，则搜索流量将带来 800 个订单。如果平均每单金额是 100 元，毛利 30 元，则搜索流量每月将贡献 8 万元销售额，2 万 4 千元毛利。

上面的计算已经非常明了。搜索订单转化率小于等于网站平均转化率。有的网站搜索订单转化率要大于平均转化率，因为搜索流量质量较高，用户有较强的购置意图。也有的网站搜索流量转化率低于平均转化率，因为回头客很多，直接点击流量转化率很高。所以如果网站有一定的累积数据，还要根据流量分析和数据做改正。

3.5.1 关键词趋势波动和预测

通过热点关键词带来搜索流量，不一定适合所有网站，前面讨论过，有效流量才是我们需要的，一个电子商务网站通过无关的社会热点带来流量价值到底有多大，要看网站营销人员是否能把自己的产品和新闻建立一定的逻辑或情感联系。如果从哪个角度看都毫无关系，这样的流量就没有什么意义了，除非网站纯粹靠广告营利。所以我们要从以下几点分析判断其趋势。

1. 长期趋势

同样是电子产品，录像机和 DVD 机都已经是昨日黄花，关注的人越来越少，而智能手机越来越受关注，这种关键词搜索长期趋势对网站主题及内容的选择，甚至产品研发、是否进入某个行业都有决定性影响。搜索次数持续下降的关键词不一定就不能做，但是企业和 SEO 人员都应该有正确的心理预测，就算排名很好，流量还是会下降。

2. 季节性波动

有很多关键词随季节正常波动，最明显的是各个节日，在节日的前后一段时间搜索量剧增，而其他时间很少有人关心。这种随季节波动比较大的关键词，SEO 人员应该事先了解趋势，提前做出内容建设、外链建设等方面的准备，有时候可能还需要开设专题，进行产品促销等。由于搜索引擎收录页面、计算排名都需要一段时间，因此 SEO 人员针对这些季节性关键词做排名也必须提前准备。

3. 社会热点预测

每一次出现社会热点新闻，都会带动一批关键词搜索次数大增。对社会新闻的敏感度是捕捉热点关键词的关键。除了关键词排行榜，电视报纸上的新闻都可能是灵感的来源。谁能先捕捉、预测到哪件事会成为热点，谁就将得到这些流量。除了针对热点组织相关内容，拥有一个圈中比较高、爬行频率也高的域名也有很大帮助，否则有了内容却不能被快速收录，也无法达到效果。

3.5.2 从赶集网广告变成“赶驴网事件”得到 SEO 启发

前一段时间，赶集网在央视的广告很疯狂，广告创意确实吸引人，不过那头驴太抢风头，结果造成了很多人都记住了驴，竟然没有记住赶集网，甚至把赶集网记成了“赶驴网”，于是“赶驴网”这个关键词火了，每天搜索量上千，甚至一度超过了一万，图 3-9 所示为“赶驴网”的百度指数。



图 3-9　关键词“赶驴网”百度指数

最早发明赶驴网这个关键词的是王建硕，他们的团队人虽然不多，但是执行力非常强。王建硕发明赶驴网这个关键词火了起来之后，立刻在百度上投放广告，购置这个关键词。同时，百姓网还注册了 ganlvwang.com，当时这个网站已经排在第一名了。后来更多的网站效仿。的要知道，搜索赶驴网的用户以前肯定不知道百姓网，本来上网找赶集的，谁知道上了百姓，那么赶集网在央视花大价钱投放的广告效果，不少就被截流到了百姓网。接着相关的新闻出来了，新闻曝光又帮百姓网做了不少广告。

于是乎，赶集网有点急了，想赶快把“赶驴网”这个词做上来，结果在自己网站首页友情连接地位第一个中，用“赶驴网”做了一个链接，这引起了李开复老师的好奇，接着，李开复老师就在微博上考同学们了，内容如下：@李开复:考考大家:赶集网首页下方有个“赶驴网”的友情链接，但是点上去才发现是链接到赶集网，又回到原页。为什么？李开复的微博粉丝有几百万，这样一来就有更多的人关注了。下面来分析，该事件对 SEO 的启发。

1. 机会

赶集网广告中的驴太吸引眼球了，于是用户把赶集网记成了赶驴网，类似的事情一直以来都不缺乏。如很多网友都会把盗梦空间，打成“盗墓空间”，会把酒店预订打成酒店预定。

每个行业中都有众多错别字，而这些错别字对 SEO 来说，非常容易，因为几乎没什么竞争，所以要多多留心，你能够发现很多机会 。

针对“赶驴网”的优化，百姓赶驴网采取的措施如下。

（1）注册拼音域名：ganlvwang.com。

（2）直接调去 baixing 网的数据生成一个新网站，并且 title 和 meta 都围绕“赶驴网”这

个词来设计。

2. 通过百姓网首页给这个新站做链接

赶集网采取的措施：仅仅在网站首页用“赶驴网”自己给自己做了一个链接。

相比之下，站在 SEO 的角度，百姓赶驴网肯定比赶集网有优势，所以在 Google 的自然排名中，百姓赶驴网排第一，赶集网排第二。

3. 百度和谷歌 SEO 策略的不同

虽然这两个网站针对“赶驴网”一词在谷歌中排名很好，但是在百度中搜索“赶驴网”，根本搜索不到这两家网站。这是为什么呢？只有一个原因：google 是纯机器来排名，百度有大批的人工参与。

百姓网为何在百度中没有排名？

（1）百姓赶驴网此举纯属投机行为，纯为 SEO 而做的，意图很明显，所以百度人工也会给其降权。

（2）百姓网针对“赶驴网”这一关键词投放有广告，如果让你排名上来了，你就停止投放这个关键词的广告了，百度岂不少赚。所以百度人工控制不让你上来。

赶集网为何在百度中没有排名？

（1）链接关系：A 网页给 B 网页做链接，那么 A 网页就是 B 网页的反向链接，这种链接有利于提高 B 网页的排名。但是赶集网目前的行为是：A 网页给 A 网页做链接，这不是反向链接，实乃多此一举的行为。

（2）赶集网不是百度的客户，百度虽然知道搜索赶驴网的都是找赶集网的，但是也不会这么轻易让你排上来。

4. 通过此时百度和谷歌的排名情况，则可以更进一步的懂得谷歌和百度，针对 SEO 策略总结如下

（1）针对 Google。关键词布局做好，反向链接多多上，排名自然不是问题。

（2）针对百度。工作要做，但是要考虑百度的人工维护机制，投机取巧太明显并且可能会让百度少赚到钱的事别做。不过，不管针对那家搜索引擎优化，它的算法都是在模拟人的思考，一切都是为了让客户有更好的体验，同时让自己合理的赚钱。考虑用户体验，同时考虑搜索引擎利益，你就不容易犯错。

3.6 金字塔形关键词分布

优化多个关键字是很多初学 SEO 的人感到困惑的问题。明显不可能把这么多关键字都放在首页上，要不然页面内容撰写、链接应用、内部链接以及锚文字的安排都将无所适从。以

权重的流动来看，可以把它比作是一颗大树。站点的主导航栏就是我们的根茎，而面包屑、分类就是树干，至于内容页面就是这颗树的枝叶。那么如何让站点权重可以合理地分配到我们根茎、树干和枝叶呢？

一个比较合理的整站关键词布局类似于金字塔形式。

- 核心关键词位于塔尖，只有两三个，使用首页优化。次一级关键词相当于塔身部分，可能有几十个，放在一级分类（或频道、栏目）首页。意义最相关的两三个关键词放在一起，成为一个一级分类的目标关键词。
- 再次一级则放置于二级分类首页。同样，每个分类首页针对两三个关键词，整个网站在这一级的目标关键词将达到几百上千个。小型网站经常用不到二级分类。
- 更多的长尾关键词处于塔底，放在具体产品（或文章、新闻、帖子）页面，如图 3-10 所示。

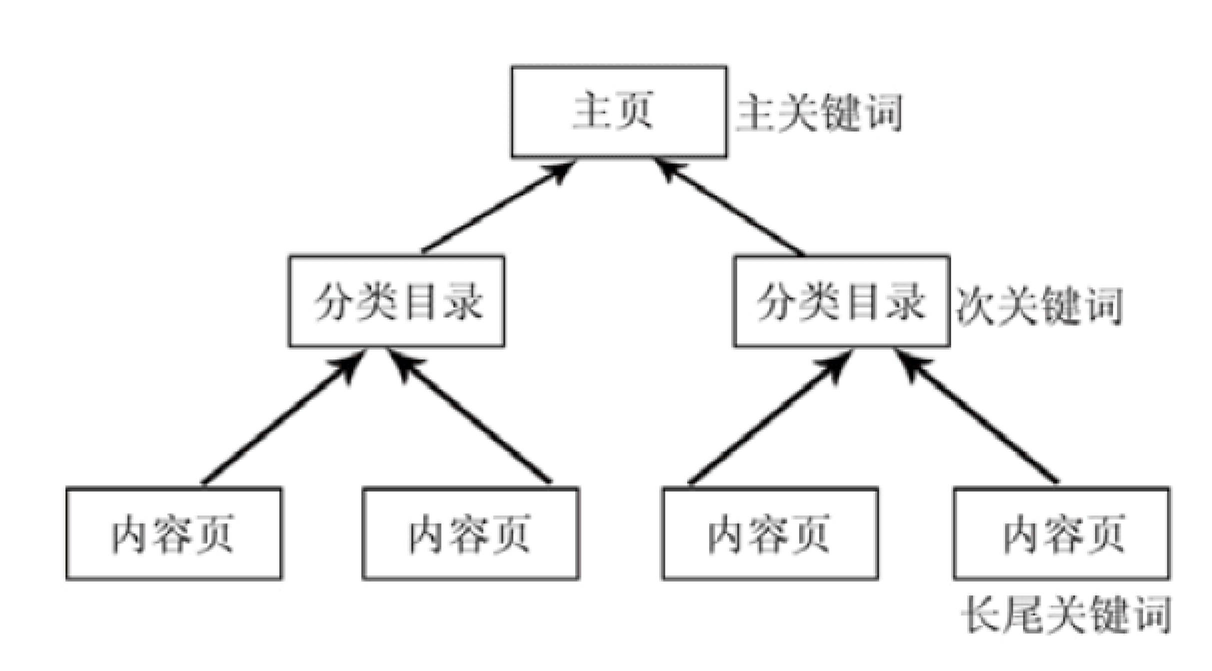


图 3-10 关键词结构分布图

得到关键词扩展列表后，重要的一步是将这些关键词有逻辑性地分组，每一组关键词针对一个分类。

假设核心关键词确定为云南旅游，次级关键词可能包括昆明旅游、丽江旅游、大理旅游、西双版纳旅游、香格里拉旅游、临沧旅游等，这些词放在一级分类首页。

每个一级分类下，还可以有二级分类。如大理旅游下又可以设置大理旅游景点、大理旅游地图、大理旅游攻略、大理美食、大理旅游交通、大理旅游自由行等，这些关键词放在二级分类首页。

再往下，凡属于大理地区内的景点介绍文章，则放在大理旅游景点二级分类下的文章页面。这样，整个网站将形成一个很有逻辑的结构，不仅用户浏览起来方便，搜索引擎也能更好地理解各个分类与页面的内容关系。

有的行业并不像旅游一样有地区这种明显的划分标准，所以关键词分组的逻辑性并不明显。比如“减肥”一词，就需要在进行关键词扩展时按行业常识将关键词分成多个组别。

经过关键词扩展得到大关键词列表后，按搜索次数排序，整体观察这些关键词可以从逻辑意义上分为：饮食减肥、减肥方法、局部减肥、快速减肥、运动减肥等一级分类。

然后将所有关键词按上述分类分组，放在不同表内可以看到，凡是和局部减肥有关的词就放在局部减肥表内，像饮食减肥、快速减肥等关键词也会做同样处理。从一级分类列表里

可以看出又能分为哪些二级分类，如局部减肥可以分为腹部减肥、大腿减肥、脸部减肥等。使用 Excel 将关键词合并、排序、分组后，整个网站的关键词金字塔结构就清晰地展现出来了，如图 3-11 所示。

序号	关键字	搜索量
1	减肥	18024
2	减肥方法	12771
3	瑜伽减肥视频	7771
4	怎样减肥	6849
5	如何减肥	5821
6	减肥药排行榜	5567
7	减肥食谱	5512
8	怎么减肥	5081
9	减肥操	3939
10	怎么减肥最快最有效	3506
11	郑多燕减肥舞	2696
12	吃什么减肥	2473
13	怎样减肥最快最有效	2399
14	减肥网	2141
15	减肥产品	2030
16	减肥吧	1740
17	减肥药	1721

图 3-11　关键词分级列表

3.7 长尾关键词

长尾理论是著名的《连线》杂志主编 Chris Anderson 在 2004 年开始在《连线》杂志发表的系列文章以及后来出版的《长尾》一书中具体阐述的。Chris Anderson 研究了亚马逊书店、Google 以及网上录像带出租网站 Netflix 等的消费数据，得出长尾理论。

3.7.1 长尾理论

长尾理论（The Long Tail）是网络时代兴起的一种新理论，由美国人克里斯・安德森提出。长尾理论认为，由于成本和效率的因素，当商品储存流通展示的场地和渠道足够宽广，商品生产成本急剧下降以至于个人都可以进行生产，并且商品的销售成本急剧降低时，几乎任何以前看似需求极低的产品，只要有人卖，就会有人买。这些需求和销量不高的产品所占

据的共同市场份额和主流产品的市场份额相比，甚至更大。

根据维基百科，长尾（The Long Tail）这一概念是由《连线》杂志主编 Chris Anderson 在 2004 年 10 月的《长尾》一书中最早提出，用来描述诸如亚马逊和 Netflix 之类网站的商业和经济模式。

“长尾”实际上是统计学中幂律（Power Laws）和帕累托分布（Pareto distributions）特征的一个口语化表达。

过去人们只能关注重要的人或重要的事，如果用正态分布曲线来描绘这些人或事，人们只能关注曲线的“头部”，而将处于曲线“尾部”、需要更多的精力和成本才能关注到的大多数人或事忽略了。如在销售产品时，厂商关注的是少数几个所谓“VIP”客户，“无暇”顾及在人数上居于大多数的普通消费者。而在网络时代，由于关注的成本大大降低，人们有可能以很低的成本关注正态分布曲线的“尾部”，关注“尾部”产生的总体效益甚至会超过“头部”。如某著名网站是世界上最大的网络广告商，它没有一个大客户，收入完全来自被其他广告商忽略的中小企业。安德森认为，网络时代是关注“长尾”、发挥“长尾”效益的时代。

举例来说，我们常用的汉字实际上不多，但因出现频次高，所以这些为数不多的汉字占据了广大的红区，如图 3-12 所示。

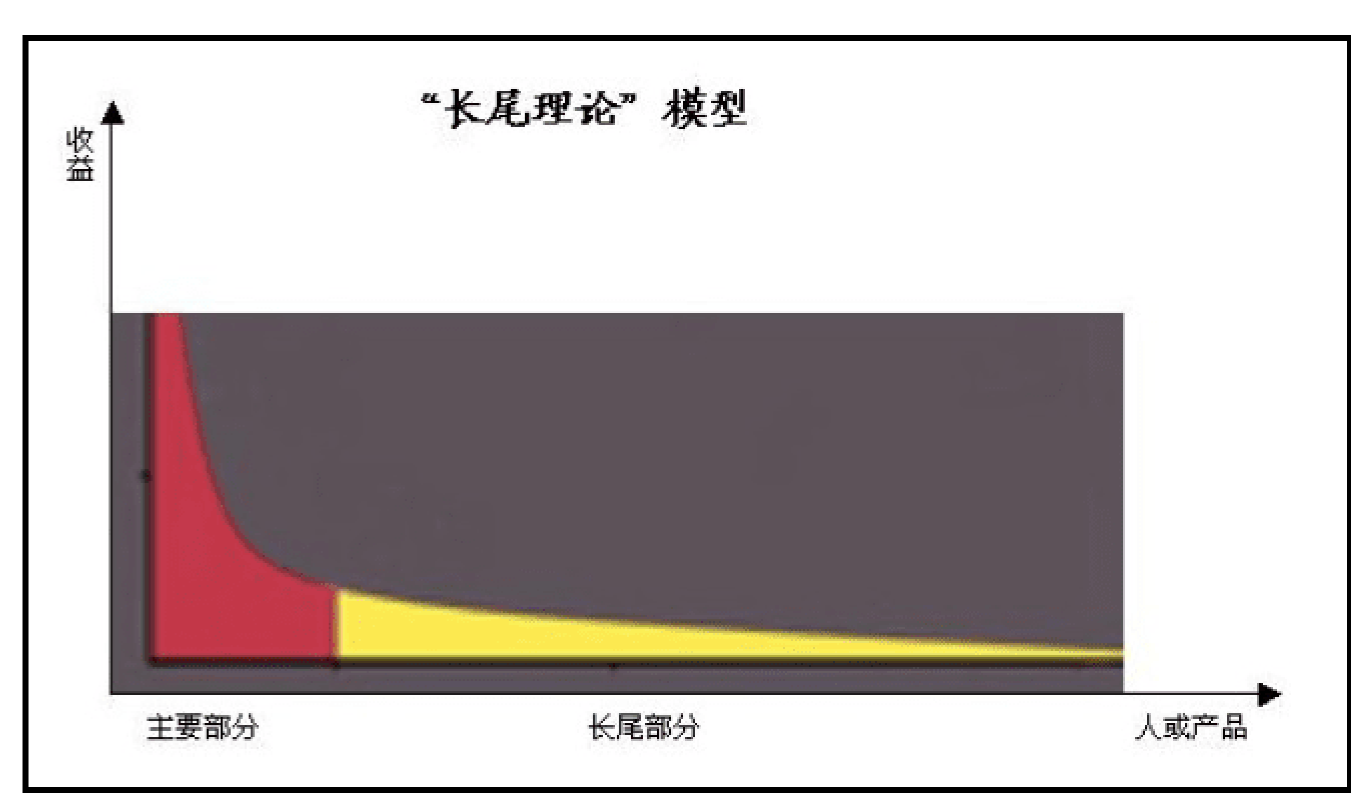


图 3-12　长尾理论

绝大部分的汉字难得一用，它们就属于长尾。Chris 认为，只要存储和流通的渠道足够大，需求不旺或销量不佳的产品共同占据的市场份额就可以和那些数量不多的热卖品所占据的市场份额相匹敌甚至更大。

长尾市场也称之为“利基市场”。“利基”一词是英文“Niche”的音译，意译为“壁龛”，有拾遗补缺或见缝插针的意思。菲利普·科特勒在《营销管理》中给利基下的定义为：利基是更窄地确定某些群体，这是一个小市场并且它的需要没有被服务好，或者说“有获取利益的基础”。

在 SEO 领域，较长的、比较具体的、搜索次数比较低的词就是长尾关键词。单个长尾关键词搜索次数小，但总体数量庞大，加起来的总搜索次数不比热门关键词搜索次数少，是流量金矿。另一个 SEO 人员关注长尾关键词的重要原因是，长尾关键词竞争小，使大规模排名

提高成为可能。

搜索领域是长尾理论最明显的体现，因为渠道足够宽，每个网民的电脑都是渠道，送货、生产成本低，搜索引擎返回每一个关键词结果的成本几乎可以忽略不计，用户需求多元化。

搜索引擎工程师确认过，被搜索的关键词中有很大一部分搜索量很小，但总体数量庞大。甚至有的搜索词以前从来没有出现过，以后也再没有出现过，只被一个用户搜索过一次。

在长尾这个词被发明以前，SEO 行业早就确立了同样的关键词原则，只不过没有长尾关键词这个名词而已。长尾理论被提出以后，最先经常使用的就是 SEO 行业，因为这个词非常形象贴切地说明了大家一直以来已经在遵循的关键词选择原则。

很多站长从流量统计中也可以明显看到长尾现象。主要热门关键词就算排名不错，带来的流量也经常比不过数量庞大的长尾关键词。可以说，长尾关键词是大中型网站的流量主力。大部分大型网站长尾流量应该占到一半以上，达到百分之七八十以上也不罕见。所以做好长尾是增加流量的关键之一。

长尾效应在小网站上较难发挥力量。大型网站主要关键词就算每天能带来几千访问量，与网站的几万甚至几十万日流量相比，还是个小小的零头。真正带来大量流量的，还是那几十万的长尾网页。小网站没有大量页面做基础，也无法有效吸引长尾搜索。

因此，在目前 SEO 过程中，无论是从网站营运积极方面还是消极方面，我们给予百度必要重视的同时，必须坚决地实施“不管这个篮子多么的牢固安全，所有鸡蛋不能都放在一个篮子里”的 SEO 策略。更何况，这样的 SEO 策略还能有效地避免“过热竞争”，给网站运营带来意收不到的收益。

除此之外，“长尾理论”在 SEO 过程中的利用还有很多，譬如“美国留学”这样的热门关键词仅竞价网站就有 6 家，但我们只要将“美国留学”分解为“美国州名+留学”，如“弗吉尼亚留学”百度排行榜上就只有一家推广网站，而且排名靠前网站资料都不完整，只要我们系统化地做好弗吉尼亚州所有对华招收留学生的大学资料，做好相关优化措施，排名就很容易上去的。而且我们只要踏踏实实做好 20 个“美国州名+留学”的相关资料，网站相关关键词排名就会上去，最后不愁“美国留学”这样的热门关键词排名上不去。

由此看来，“长尾关键词”只是“长尾理论”的一种 SEO 应用手段，掌握了它并不等于我们完全掌握了“长尾理论”，只有熟练地掌握并活学活用长尾理论后，对于我们拓宽 SEO 工作思路，做好 SEO 工作，有效化解压力，提高工作效率大有裨益。

“长尾理论”对网站营销推广运用非常重要——它并不仅仅局限于 SEO，如“新闻营销”、“差异化营销”、“邮件营销”、“短信营销”、“微博营销”等方法，这些方法在网站营销上特定领域都具有不可比拟的优势，但都存在着自己的局限性，只有与时俱进，应用好各种方法，才能真正把网站营销工作做好。

最后，必须强调的是，我们如此对待世界搜索引擎规则制定者谷歌太过于“功利性”的“看人下菜碟”态度表明，一直强调“用户体验”的我们并没有真正地理解“长尾理论”，持这样的态度，我们不可能真正地做好网站服务，使网站达到良好的“用户体验”的。

“长尾理论”必须建立在对每一个“小众用户”的充分尊重与理解的基础之上，尽自己最大所能满足对方的需求，只有这样，才能真正做好网站。因此，我们在网站营运中一方面要把中国人那种“一个人都不得罪”、“多条朋友多条路”的处世观发挥到极致；另一方面，要真正地从心里尊重每一个人（无论是用户，还是服务提供商），不能把任何人看作是“有他不多，无它不少”的“牛脖子上的耷拉皮”，才可能真正做好网站营销推广工作。

这些质疑道出了中国所有站长（SEOer）的心声（虽然更多人不会这样直白）。中国大陆网站谷歌搜索流量是“王小二过年，一年不如一年”，如现在 gouyn12 服务网站谷歌收录 12 万条，可一个月从谷歌来的搜索流量寥寥无几。

但我们用如此的态度对待谷歌是错误的，这表明了我们对“不管这个篮子多么的牢固安全，所有鸡蛋不能都放在一个篮子里”的“长尾理论”的通俗化表达并没有真正理解。

虽然我们对长尾关键词的定义、做法说起来都头头是道，但令人遗憾的是，“不管这个篮子多么的牢固安全，所有鸡蛋不能都放在一个篮子里”就是长尾理论的通俗表述——只不过是较为消极保守的表述。长尾关键词就是基于不管这个关键词多么的强势稳固，发展态势良好，但为了最大化的网站利益，也不能将所有网站的流量都寄托在这个强势关键词上，必须深入挖掘更多并不强势稳固的关键词，并把它们做好，才能真正将网站 SEO 工作做好，达到提高网站转化率的最终目的。

显然，这样的长尾理论思维也必须表现在站长对待搜索引擎的态度上。的确，现在对于中国大陆网站来说，百度强势，关注百度搜索引擎优化是每个站长的必然选择，但如果我们把注重百度搜索引擎的精力分 10%给其他“长尾”搜索引擎，可能就会有出乎意料的惊喜。

3.7.2 怎样做长尾关键词

做好长尾既简单又困难。说它简单是因为，一般来说不需要也无法做深入关键词研究，也不需要刻意优化特定长尾关键词。由于数量庞大，去查看搜索次数、专门调整页面优化都是不可行的，只能通过大量有效内容及网站结构方面的优化确保页面收录。只要页面基本优化做好，长尾关键词排名就能全面提高。

有的站长在论坛中询问，怎样做长尾关键词的研究？其实这很难研究。一个网站做几百几千个关键词的研究有可行性，但这个数量算不上是长尾词。真正体现长尾效应的网站至少要几万个页面，达到几十万数百万也只是普通的。长尾关键词数量至少上万。大致列出这些关键词，通过软件查询搜索次数或百度指数是可行的，但再进一步研究，如估计竞争程度、分配具体页面等，可行性及必要性大大降低。

所以做好长尾词的关键在于收录和页面基本优化，这两方面都是网站整体优化时必须要做的。不必考虑特定关键词，但结果是长尾词会全面上升。

说它难在于，做好长尾首先要有大量内容，对中小企业和个人站长来说，除了转载、采集，似乎没有更好的方法，除非网站是用户产生内容。

网站基本优化，尤其是内部链接结构，必须过关，才能保证大量包含长尾关键词的页面被收录。对一些大型网站来说，保证收录并不是一件简单的事。

要提高长尾词的排名，域名权重也是个因素。网上相同或相似内容很多，域名权重低，页面排名必然靠后。

其次改进长尾关键词。现在不但是主关键词的竞争非常激烈，就连一些相关性极度强的长尾关键词也是竞争非常激烈的。就像藏獒这个关键词扩展出来的长尾关键词，如图 3-13 所示。

请输入你要查询的关键词：藏獒 查 询

序号	关键字	搜索量	收录数	首页第1位	首页第2位	优化难度
1	藏獒	15953	25000000	www.315hn.com 百纳藏獒网—中国藏獒在线网	baike.baidu.com 藏獒_百度百科	
2	藏獒价格	2618	2030000	www.315hn.com 百纳藏獒网—中国藏獒在线网	www.58.com 【藏獒】藏獒价格藏獒图片-	
3	藏獒图片	1932	2270000	image.baidu.com 藏獒图片_百度图片	www.315hn.com 藏獒鉴赏_纯种藏獒图片大全	
4	藏獒打架视频	1614	835000	video.baidu.com 藏獒打架视频_百度视频	v.youku.com 藏獒打架火爆激烈-视频-优酷	
5	藏獒多吉	1180	737000	www.iqiyi.com 《藏獒多吉》动漫_高清在线	baike.baidu.com 藏獒多吉_百度百科	
6	藏獒视频	1127	1610000	video.baidu.com 藏獒视频_百度视频	v.youku.com 藏獒打架火爆激烈-视频-优酷	
7	藏獒在线	1025	1510000	www.aiao.cn 藏獒在线—中国藏獒专业权威	bbs.aiao.cn 藏獒在线—中国藏獒最大专业	

图 3-13　藏獒长尾关键词挖掘表

从查询工具中可以查到，其中藏獒这个关键词扩展出来的前六名的长尾关键词其竞争度都是非常激烈的，而竞争激烈了自然优化难度也随之上升了。如果你硬是要优化跟竞争度跟主关键词差不多的，那需要的时间就非常之久了。而要想优化快速取得效果，那么就需要稍微改进这些长尾关键词。如我们是买卖藏獒的必然需要藏獒价格这个词，那我们可以做纯种藏獒价格、藏獒幼犬的价格、北京藏獒价格等，虽然搜索量不如主要关键词，但是长尾词数量多了之后它们带来的流量一样不可小视。

尽管现在的网站行业的竞争非常的激烈，但是每个市场都会有一些存在于角落的小市场还没被发掘。如现在比较火热的软文推广，许多的站长都知道软文可以增加外链、提高曝光率和增加额外流量，但是软文就只有这么几点吗？其实不然，软文在不同的人群中发挥的效果和作用都是不相同的，如淘宝的就是提高转化率，名人的就是提升知名度，个人站长的就是增加外链。同样的在咱们优化了一定量的长尾关键词获取好排名之后，不应该满足于现在的状况，应该试着创造属于自己的长尾关键词。如选择一个比较冷门的长尾关键词，把它打造成自己站点的一个品牌长尾关键词，就像搜索新闻时，大部分用户都会这样搜索——新浪实时新闻、新浪娱乐新闻、网易新闻、网易魔兽世界等。在笔者看来创造长尾关键词比优化一个竞争激烈的关键词容易得多，同时自己创造出来的长尾关键词更利于掌握其效果，而且

优化推广起来也是非常容易的。创造长尾关键词对于应对竞争越发竞争的市场可谓是一条不可多得的好途径。

很多站长对于以长尾关键词作站点的主要优化对象抱以嗤之以鼻的态度，认为优化流量小的长尾关键词，何年何月才能让自己的站点流量上升到某种高度啊。笔者认为这样的站长其实就是一些眼高手低的，试想一下，如果连冷门的长尾关键词都认真对待，怎么可能面对长期无排名无效果的热门关键词呢？而且优化长尾关键词还有一个好处，那就是可以以量取胜。

所以说，站长在优化关键词时，选择长尾关键词确实是一条不错的途径。长尾理论是SEO 人员必须理解和关注的概念，但是要真正显示长尾关键词效果，却不能从关键词本身出发，而是从网站架构、内容及整体权重上着力。预估流量及价值正规公司关键词研究的最后一步是预估搜索流量及价值。

个人站长做关键词研究不一定需要这一步。找到最合适的关键词就可以直接去做，尽力能做多少就做多少。但正规公司，尤其是大公司则不行。整个 SEO 项目是否能获得批准，又能否获得公司高层支持，申请预算、安排人员、工作流程及时间表等，都需要 SEO 人员提供明确的预计搜索流量及给公司带来的价值。

3.7.3 关键词组合

网站上的内容与关键字相互对应而成，从某种意义上来说，搜索引擎优化就是建立在页面内容基础上的。当我们确定了网站的核心关键词以后，就需要辅以大量的相关内容来支撑该核心关键词或其他的相关长尾关键词，以让搜索引擎得知这个网站是做什么的，这个网站的核心是什么。那么这些长尾关键词该如何组织起来呢？我们就从以下八个方面来说说网站SEO 中的关键词组合方法以及如何通过关键词组合来定制和更新网站内容。

1. 第一策略：产品或服务+特点功能（质地、类别、行业背景等）

这一类内容是产品和服务的底层，但是也可以延伸出去。可以是对自身产品的介绍和功能阐述，也可以是该产品的原料产地、背景知识等。如果是某服务，可以对该行业专有名词进行解释，介绍行业新闻，技术文章以及服务流程等。

如 SEO 大赛关键词为：棉鞋。你要做一个棉鞋的企业网站，那么网站内容可从产品自身特点来展开，如棉鞋的质地、棉鞋制作工序、棉鞋样式和棉鞋花纹等方面。也可以从人群分类来展开，如儿童棉鞋、男士棉鞋、女士棉鞋、休闲棉鞋等。同时还可以追根溯源，如 xx 的起源、xx 的原产地等。

2. 第二策略：产品或服务+搜索意图（是什么，怎么用，哪里有卖等）

隔行如隔山，每个行业都有自己特有的规律。了解和熟悉这个产品或服务所在行业的规则，也就会了解很多特有的关键词组合方法。

如冬虫夏草是什么？某美白产品怎么用？某车膜哪里有卖等。产品和搜索意图的关键词

组合，搜索这些词的用户都是网站的潜在用户，转化为付费用户的几率也非常大。

对于这类关键词，大家可以到维基网站、问答网站参考。往往用户提交的咨询问题，就是很好的关键词。

这里有一个小建议：不同阶层的网民搜索的关键词也有很大不同。不过大多数网名搜索的关键词都比较低端，偏向于口语化。在网站建设中有一个说法就是做网站，如有一个关键词：北京做网站哪家好？我们其实可以为这个关键词来组织一篇文章，标题就叫：北京做网站哪家好。

3. 第三策略：产品或服务+品牌型号（品牌名称、通用型号等）

这里的品牌可以是自己的品牌，也可以是其他甚至是竞争对手的品牌；可以是知名品牌，也可以是非知名品牌；这里的型号是指行业的标准，产品的特有型号和规格的技术参数等。以上只是一个例子，但是不限于这些。

如接过轴承相关单子的朋友可能比较熟悉，轴承的关键词很多就是通过不同品牌，不同型号来组合的。如深沟球轴承、ASK NUE70 轴承等。把这个思路用在其他的产品和服务中，我想效果也是不错的。

世界上十大打火机品牌，你能说出几个？大家一定会想到 Zippo，没错，Zippo 就是十大打火机品牌之首。

行业先驱们打造出的这些品牌我们很难超越，业内的行业规范我们很难改变，但是这些耳熟能详的产品所带来的人气是巨大的，我们都可以以不同的形式融进网站的内容中，从而带动网站在搜索引擎上的排名提升。

4. 第四策略：产品或服务+经营模式（批发、代理、加盟等）

把经营方式融入网站内容。在发布产品信息时，我们希望更准确地找到客户，也希望找到更专业的客户。这时，我们就需要在信息中标明自己想要达到的目的。

如我们想找到代理加盟商，就可以在信息名称中加“代理加盟”作为信息的主关键词；如果想做大量的批发生意，就可以把“批发”作为信息的主关键词。这里的信息可以是标题，可以是网站的部分内容，也可以是网站的其他地方。

结合实体的产品来看就是：加入资讯或文章模块，引入淘宝和独立网站代发货的模式。

现在越来越多的淘宝用户苦于找不到合适的货源，而也有非常多的货源提供商找不到下一级的代理。也就是说，如果在通域商城相关页面增加代理模块，那么他可以设置大量的类似如下关键词：诺基亚 N95 代发货、phone 代发货。

5. 第五策略：产品或服务+服务方式（培训、函授等）

服务方式边缘元素。网站 SEO 中的八大关键词组合策略，如 SEO 优化、SEO 搜索引擎优化、SEO 关键词优化，SEO 优化教程的不同等，面向的潜在用户人群也不一样。

如某某考试培训、某大学函授课程、某化妆品邮购，还有目前比较流行的 QQ 群培训，实体的面授培训等。

6. 第六策略：产品或服务+企业信息（公司、快递等）

这在传统的服务性企业中比较常见，特别是一些便民服务行业。

如深圳货运公司电话、中关村搬家公司、中关村送水公司、石家庄房屋中介公司电话等。还有一些如某产品厂家地址、Pagebbc 官网地址等。

7. 第七策略：产品或服务+错字别字（易混淆品牌和名称等）

很多朋友把这一点叫做另类 SEO，其实并非如此，互联网的内容丰富和开放造就了丰富多彩的错别字现象。此前在 QQ 群中广泛流传的“建国大业首映没”表情图片事例相信大家现在都记忆犹新。

以前看到过一个站名字叫讯雷，并且这个关键词排名仅次于迅雷官网，可想而知这个站截取到不少的流量。诸如此类还有很多例子，如刘亦非与刘亦菲、张钰与张珏、vagaa 与 wagaa、易建联与易健联等。当然这里只是举一个例子，如果你能寻找到产品与服务中类似的关键词，相信效果会非常明显。

8. 第八策略：产品或服务+领域区域（应用领域、地域名称等）

我们经常会说道 360 行，行行出状元，这里的 360 行只是泛指行业之多。同时中国地大物博，30 余个省市和地区，3000 多个县级市。笔者曾经把这 3000 多个地区名称+产品关键词用到一个客户网站上，效果是非常明显的（这些地区的关键词表大家可以到优友网论坛资源区下载）。地区名称和网站产品或服务关键词组合起来，覆盖的受众就非常之多。这一招其实已经并不新鲜，早就被大的网站用到网站和内容中了。

此类案例很容易在分类信息网站看到。如有时候在 58 同城上看到某地区关键词排名非常好，但该频道本身基本上是没有内容的。很多大的网站都是海量的关键词+模板，他们往往不在于只做某一个关键词，而是做某一批关键词。58 同城通常这样玩：北京+租房、天津+租房及非常小的县级市+租房。讲这个的目的并不是一定让大家也来玩海量关键词，而是要有一个思路，你要融入地区类关键词+产品或服务。

3.8 小结

选好关键词，最后还应该参考竞争对手是否是品牌类的网站或者企业，如果竞争对手过于强大，笔者建议最好是另设关键词或者使用长尾理论重新选择比较适合自己的关键词。怎样才算是竞争度大呢？你在搜索引擎中输入一个词，如果首页网站的网址全部是一级域名，或者前两页是一级域名，就说明这个词的竞争度很大。反之，如果首页网址是三四级就说明竞争度比较低，容易把排名做上去。另外还可以参考一些具有高权重的网站、博客以及有威望的个人，看看这些站长是如何吸收到流量的，如何进行关键词设置的，便于自己做决择。

第 4 章　URL 网页地址优化

如今随着互联网的盛行，在网站建设、网站优化中出现了很多的问题，如建网站不规范或者后期的网站不规范，最为严重的可能是导致网页 URL 标准化问题。本章重点介绍如何优化网页 URL 的问题，介绍的内容包括 URL 网页地址优化的基础知识，网页地址的命名方法，分隔符的使用，网页的地址长度定义，以及网页地址重定向和静态化的问题。

4.1　URL 简介

URL 相当于页面的地址，也称统一资源定位符（URL，英语 Uniform Resource Locator 的缩写），是因特网上标准的资源的地址。它最初是由蒂姆•伯纳斯-李发明用来作为万维网的地址的。用户或搜索引擎都必须通过 URL 才能访问相应的页面，所以对网站内部的优化也是从 URL 开始的。

在因特网的历史上，统一资源定位符的发明是一个非常基础的步骤。统一资源定位符的语法是一般的，可扩展的，它使用 ASCII 代码的一部分来表示因特网的地址。一般统一资源定位符的开始标志着一个计算机网络所使用的网络协议。

统一资源定位符是统一资源标志符的一个下种。统一资源标志符确定一个资源，而统一资源定位符不但确定一个资源，而且还表示出它在哪里，如图 4-1 所示。

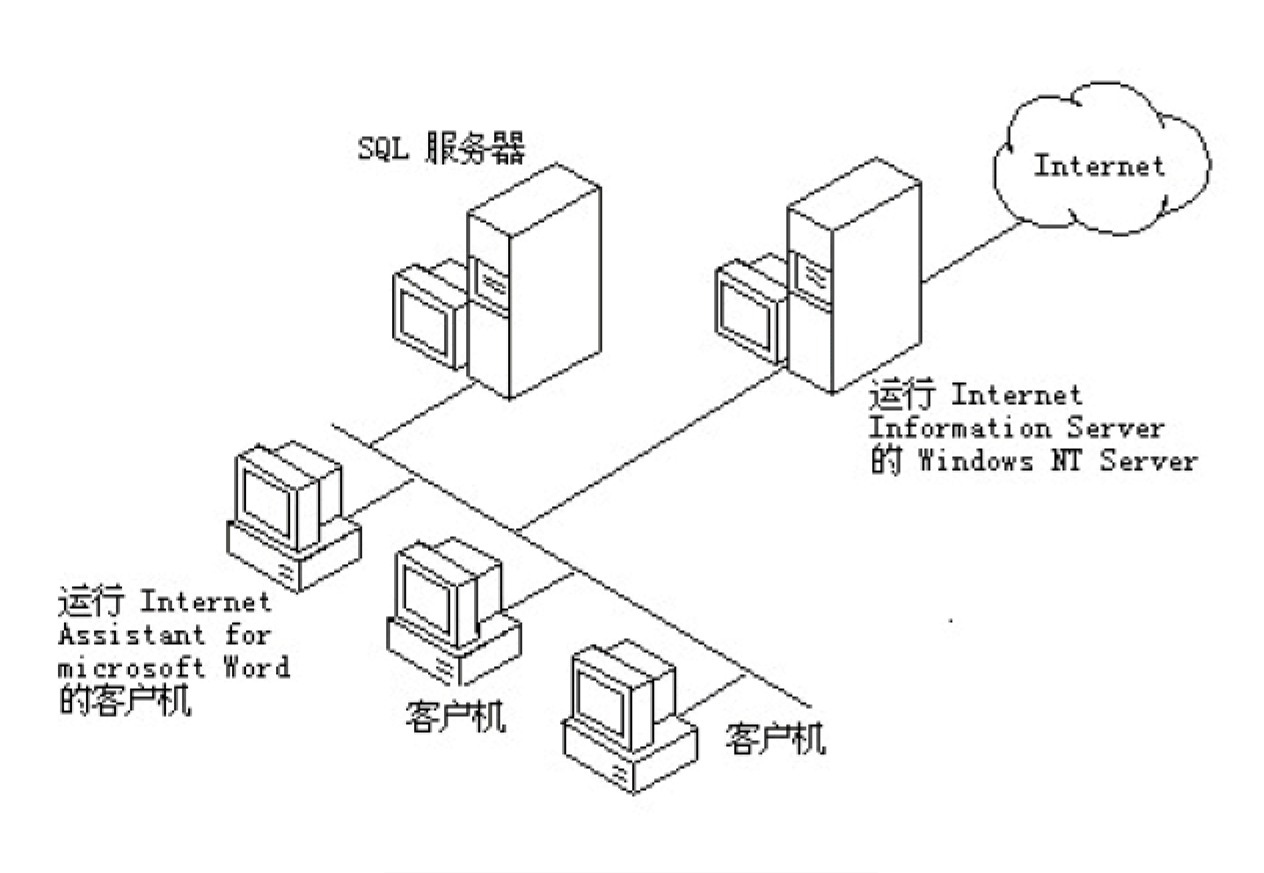


图 4-1　统一资源定位符

4.1.1　定位符语法

```
协议://授权/路径？查询
```

授权部分一般是服务器的名称或 IP 地址，有时后面还跟一个冒号和一个端口号。它也可以包含接触服务器必须的用户名称和密码。路径部分包含等级结构的路径定义，一般来说不同部分之间以斜线（/）分隔。询问部分一般用来传送对服务器上的数据库进行动态询问时所需要的参数。

完整的、带有授权部分的普通统一资源标志符语法为：

```
协议://用户名@密码:子域名.域名.顶级域名:端口号/目录/文件名.文件后缀?参数=值#标志。
```

统一资源标志符参考指的是单个的（如超文本传输协议文件中的）统一资源标志符。统一资源标志符参考分绝对参考和相对参考。上面所描写的都是绝对的统一资源标志符参考，相对参考只包括体制特殊的部分，它参考的对象位于包含这个参考的文件的一个相对位置上。统一资源标志符参考还可以由一个统一资源标志符加上一个“#”号再加上上述的统一资源标志符内的一个标志点。这个标志点不是统一资源标志符的一部分，而是让用户浏览器在获得了文件后来导航用的，因此它实际上不被送到服务器。

4.1.2　大小写

统一资源定位符一般是分大小写的，不过服务器管理员可以确定在回复询问时大小写是否被区分。有些服务器在收到不同大小写询问时的回复是相同的。地址结尾的“.”号，在互联网的发展初期，访问一个网站不是单纯的输入这样 DNS 服务器才能够识别。后来，微软公司在 Windows NT 3.51 中对其进行了修改，可以自动在 DNS 查询时自动增加一个.号，随后 UNIX，NetWare 也随之而跟进，让服务器可以识别结尾没有“.”的域名。但是，符号“.”在现在的网址中仍然可以使用，统一资源定位符的日常使用超文本传输协议统一资源定位符将从因特网获取信息的四个基本元素包括在一个简单的地址中：传送协议、服务器、端口号、路径。

4.1.3　URL

统一资源定位符 URL 是对可以从因特网上得到的资源的位置和访问方法的一种简洁的表示方法。

URL 给资源的位置提供一种抽象的识别方法，并用这种方法给资源定位。

只要能够对资源定位，系统就可以对资源进行各种操作，如存取、更新、替换和查找其属性。

URL 相当于一个文件名在网络范围的扩展。因此 URL 是与因特网相连的机器上的任何可访问对象的一个指针。

URL 的一般形式是：

```
<URL 的访问方式>://<主机>:<端口>/<路径>
```

URL 的访问方式有如下几种。

- ftp：文件传送协议 FTP
- http：超文本传送协议 HTTP
- News：USENET 新闻

其中，<主机> 是存放资源的主机在因特网中的域名。

4.2 URL 优化方式

URL 优化形式大体分为三种：静态形式、动态形式、伪静态形式。就是指通过对 URL 各组成部分进行恰当的调剂，以提高 URL 的搜索引擎友好性。第一，对域名、目录、文件的命名；第二，分隔符的使用；第三，URL 长度及关键字词频的控持。这三部分间是彼此制约、互相影响的，在优化的进程中我们必须控制这三部分的内在接洽，对各部分进行充足的和谐，才能到达最佳的优化效果，否则就功亏一篑。

URL 优化是网站优化的六大重要环节之一，占据着重要的位置。不仅在决定页面相关性方面起到重要的作用，更重要的是 URL 还会直接影响搜索引擎对页面的收录。

4.3 网页地址命名

对于 URL 的命名是针对 URL 每个组成部分（域名、目录、文件名）而进行的，也是 URL 优化过程中最重要的环节之一。网页 URL 命名的关键在于使用合适的关键词为 URL 各组成部分进行命名，意思就是说把文章关键词结合到 URL 中显示出来，这样有利于提高页面的相关性，突出页面主题的作用。需要注意的是，URL 的命名是在伪静态的情况下进行的深度优化。

通常在给 URL 进行命名的时候选择中文、英文及拼音这三种形式的关键词对 URL 进行命名，下面将具体分析这三种常用的 URL 命名形式的相关技巧。

4.3.1　中文命名

首先，如果要想使用中文形式对网页 URL 组成部分进行关键词命名，那么首先要确定你所使用的 Web 服务器必须支持以中文名称访问相应的目录或者文件（一般情况下，需要对 Web 服务器进行必要的配置后才能支持以中文名称访问相应的目录或者文件）。

在 URL 中，以关键字的中文形式为 URL 组成部分进行命名要比使用英文或拼音更简单。因为，中文形式的关键字不需要考虑分隔符等问题，只须选择与目录或者页面内容相关的中文关键字即可。但并不是所有的搜索引擎都检索含中文字符的 URL。

如百度，大家在百度进行中文搜索的时候在链接中反映出来的都是经过转码以符号的形式表现出来的，而且在搜索结果中，可能会看到在百度快照前面的绿色地址中含有中文，但是大家试试把鼠标移到标题上看看，反映出来的链接同样是转码符号。其实搜索引擎也不是很重视以中文关键字命名的 URL，笔者认为直接用汉语拼音效果可能更好，百度对拼音的识别度是很高的。

4.3.2　英文命名

将关键词转化成英文的形式来对 URL 进行命名的方式是目前最常用的 URL 命名的方式之一。但是在使用的时候我们必须要注意以下两点。

- 必须使用一个真实存在的英文单词或者一对词组来进行 URL 的命名。
- 组成 URL 的单词之间必须用分隔符进行分隔，因为将英文单词连在一起，搜索引擎是读不懂的，如 http://www.123pet.com.cn/cats/。

4.3.3　拼音命名

目前，绝大多数中文搜索引擎都支持拼音形式的关键字搜索。也就是说，中文搜索引擎可以识别拼音形式的关键字。因此，也可以利用关键字的拼音形式对 URL 各组成部分进行命名，这对提高中文页面的相关性方面起着非常重要的作用。

如在一个目录下存放与“哈士奇”相关的页面，可以以“哈士奇”的拼音形式对该目录进行命名，即 http://www.123pet.com.cn/gou/hashiqi/。

在使用关键字的拼音形式对 URL 各组成部分进行命名时，需要注意以下几点。

（1）每个字的拼音不需要采用分隔符进行分隔，搜索引擎也能识别，如图 4-2 所示。



图 4-2　“chongwuwang”搜索结果

即使使用空格进行分隔后，显示的结果也是“宠物网”，如图 4-3 所示。



图 4-3 “chong wu wang”搜索结果

（2）考察搜索引擎是否能够识别我们所选择的关键字的拼音形式。首先，在搜索引擎搜索该关键字的拼音形式，然后查看返回结果是否与我们选择的关键字一致。

在 URL 优化中，应该根据的实际情况选择其中一种最合适的关键字命名形式。同时要避免在 URL 中使用与 URL 所指向的页面主题不相关的关键字。

4.4 分隔符

为了让搜索引擎能够正确识别以英文形式命名的 URL 中的关键字，需要使用相应的符号对词组之间的单词进行分隔。

4.4.1 分隔符的使用

常见的分隔符包括：空格“　”、横杠“-”、下划线“_”、逗号“，”及加号“+”等。如搜索引擎并不能识别“digitalcamerabattery”，因为这个单词在搜索引擎的切词系统中是不存在的，但却可以识别“digital camera battery”或者“digital-camera-battery”。

说明除了使用特定的符号作为分隔符外，还可以使用其他的方式来表达单词间的分隔意义，如将每个单词的第一个字母设为大写（CamcorderBattery）。但这种方式只方便普通用户识别，对于搜索引擎来说却是无效的，因为搜索引擎并不区分单词间字母的大小写。

在 URL 中，很多符号都具有分隔符的意义。但每个符号在 URL 中所起的作用及表达的意义是不一样的。不仅不同的分隔符在 URL 中所起的作用不一样，即使是同一分隔符在不同的搜索引擎里表达的意义也会有所差别。

在有些搜索引擎里，把 URL 中的横杠“-”与空格符“%20”都当作空格符处理（URL 中“%20”就是空格）。如在 Google 中，以下关系是成立的：

camcorder-battery == camcorder battery == camcorder%20battery

下面来进行验证。以“camcorder battery”为例，在 Google 中搜索“allinURL: camcorder battery”，可以看到以加号“+”、横杠“-”及空格符“%20”作为 URL 中单词间分隔符的搜索结果。

搜索结果中的“camcorder-battery”与第二条搜索结果中的“camcorder%20battery”在 Google 中表达的意义是相同的。

接下来将重点介绍常见的分隔符：横杠“-”、空格“ ”及斜杠“/”。

（1）横杠“-”

对于分隔符的使用，可以在域名的命名规则上得到启发。如在申请一个域名时，只能使用横杠“-”而不是下划线“_”或者空格“ ”作为分隔符。由此，可以推测对于搜索引擎来说，横杠“-”比空格“ ”及下划线“_”更加友好。

（2）空格“ ”

不管是在浏览器还是搜索引擎中，URL 中“%20”字符表达的就是空格的意义，但这不利于普通用户对 URL 的理解及记忆。用户不喜欢的东西，搜索引擎也不会喜欢的。

（3）斜杠

斜杠“/”是 URL 中的自然分隔符。但是，斜杠只能作为 URL 中目录之间或者目录与文件名之间的分隔符，而不能作为目录名称或者文件名称内部的分隔符。

4.4.2 分隔符的选择

SEO 该用那种分隔符好呢，可以去看看一些大型门户站的标题分隔符是怎么用的（如百度和新浪）。

常用的分隔符有：“-”、“_”、“|”、“,”、“（）”还有空格，其中最常用的就是“-”和“_”。

“–”表示相互连接的短语有彼此解释说明的作用。如果将数个隐性语义相近的词或短语用“–”连接在一起构成一个语义更加明确的网页标题，则会在搜索引擎中获得较高的权重。

“-”分隔符与两边的短语之间有没有空格，没有特殊的规定，通常情况下是在有空余字节的时候可以加上空格。

“_”分隔符用于表示一种涵义上的从属关系。“_”前面的网页标题短语通常表示的内容更具体，常常是从属于其后面的标题短语。如：博客圈_新浪博客_新浪网。“_”分隔符两边没有必要加空格。

“|”分隔符出现在网站的标题中用于分割一个网站相对独立的数个栏目或是单独一个网页上的不同内容。各个分隔之间是有一个半角的空格，但是现在也有人不用了。

总体来讲，百度喜欢符合中国人使用习惯的“_”和“|”。谷歌是而“-”和空格。

实质上，搜索引擎并没有对网站标题关键词分隔符作出明确的规定必须用哪一种，不能用哪一种，只要合理就没有什么问题。当然我们必须要站在用户体验的角度上去想问题，所以自然而然的，就应该明白应该用哪些分隔符对网站标题关键词进行分隔。选择标题分隔符的时候可以看看大型门户是怎么做的，看看百度、谷歌、雅虎用的是什么分隔符，我们就用哪个分隔符。

4.5 网页地址长度

搜索引擎在抓取页面时，对页面的 URL 长度存在一定的限制。对于超过限定长度的 URL 指向的页面，搜索引擎就可能放弃收录。而决定 URL 长度的主要因素包括“域名长度”、“路径长度”及“文件名长度”。

1. 域名长度

域名长度是指“子域名+域名名称+域名类型”所占用的字符数。如：www.123pet.com.cn 的子域名是 www，域名名称是 123pet，而域名类型 com.cn，则域名长度为 17。

2. 路径长度

路径长度指文件存放路径的名称所占用的字符数，路径长度=目录 1 名称长度+目录 2 名称长度+……+目录 N 名称长度，如：http://www.123pet.com.cn/gou/jinmaoquan.html。其中，“/”为根目录、“gou/”二级目录，则该页面的路径长度=1+4，即 5。

3. 文件名长度

指文件名称所占用的字符数（包括后缀名），例如：“jinmaoquan.html”，文件名长度就是 15。

综上所述，URL 长度=Internet 资源类型+域名长度+端口号+路径长度+文件名长度。如：http://www.123pet.com.cn/gou/jinmaoquan.html 的 URL 长度=17+5+15，即 47。

搜索引擎抓取页面时，对页面的 URL 长度是有一定限制的，对于超过这个限制的页面搜索引擎就可能会放弃收录；而且，页面的 URL 越短，得到的权重就越高。

例如：

URL A： http://www.123pet.com.cn/zangao.html

URL B： http://www.123pet.com.cn/gou/zangao.html

同样的页面 zangao.html，使用 URL A 比 URL B 能获得更高的权重。

4.6 URL 组成部分优化

域名是一个网站的入口，是所有用户（包括搜索引擎）访问网站的必经之路。在某些情况下，域名还能从一定程度上反映网站的主题。如网站 123pet.com.cn/的主题是 pet（宠物）。

1. 域名的组成

以 123pet.com.cn/这个域名为例，它是由两个不同的部分组成的，这两个部分间具有鲜明的层次关系。其中，com.cn 是域名的第一层，即域名类型；而 123pet 则是域名的名称，处在第二层。

2. 域名注册时间

域名注册时间就是域名建立的时间，域名注册的时间越长，从一定程度上反映这个域名是更值得信任的。因此，这个域名所指向的网站就可以获得更高的权重。下面介绍怎样查看网页注册的时间。

在实际应用中，我们会使用域名信息系统查询域名的注册时间、过期时间及域名所有者等信息，每一个域名注册提供商都会提供这样的免费查询服务。如我们在注册某个域名的时候就会利用域名信息查询系统查询相关域名的信息。

3. 域名使用时间

域名使用时间是指域名被绑定到某个特定的网站以后，这个网站运营的时间。但是，域名的使用时间需要结合网站的主题进行衡量。也就是说，域名的使用时间是与域名所指向的网站的主题相关联。如一个域名使用了 10 年，如果这个域名在这 10 年间，断断续续地指向 100 个不同主题的网站，那么这个域名在提高网站主题相关性方面所起的作用也是微不足道的。

4. 域名过期时间

每一个域名都有一个使用时间，这个使用时间主要是取决于购买或者注册域名时所付出的费用（可以购买三年甚至是更长的使用时间）。

站在搜索引擎优化的角度，域名的过期时间也会在一定程度上影响网站的权重。一个域名的可使用时间越长，那么得到的权重相对来说就会越高。如对于同样时间注册的两个域名，域名 A 的过期时间是 10 年后，而域名 B 的过期时间是一年后，那么搜索引擎就会认为域名 A 更值得信赖，从而赋予域名 A 更高的权重。

5. 关键字域名

使用关键字作为域名，还涉及关键字的表达形式问题，例如中文、英文及拼音等。在实际使用中，我们应该根据网站的具体情况采用相应的命名形式。如对于简体中文网站，我们常见关键字的拼音形式作为域名。

来看看下面这几个关于关键字域名的例子。

域名（1）：www.shumaxiangji.com

域名（2）：www.smxj.com

域名（3）：www.digital-camera.com

这 3 个域名都是想在域名中表达“数码相机”这个关键字，却存在很大的差异。

域名（1）中 shumaxiangji 是关键字“数码相机”的拼音形式，如果这是一个与“数码相机”相关的简体中文网站的域名，则能很大程度上提高网站的相关性。

域名（2）中的 smxj 是“数码相机”每个拼音的首字母，但搜索引擎的切词系统里并不存在这个词。因此，该域名也就表达不了“数码相机”的意义。

域名（3）中的 digital-camera 是“数码相机”的英文形式，绝大部分搜索引擎都能识别，因为 digital camera 是一个存在于搜索引擎切词系统里的词组。对于选择该关键字作为域名的网站，如果内容是英文，则该域名的作用是非常大的。但如果内容是中文，则不如域名（1）。

6. 域名后缀问题

常见的域名后缀包括：.com、.net、.org、.gov 和.edu。其中，.com 表示商业机构，.net 表示网络服务机构，.org 表示非营利机构，.gov 表示政府机构，.edu 表示教育机构。

实践证明，Google 更重视.org、.edu 为后缀的域名，这说明 Google 更重视非营利性与教育性网站。但是，我们不能因为搜索引擎特别重视这两类域名而忽略网站自身的性质，采用不相关的域名后缀。

域名后缀只是决定网站排名的众多因素之一，并非唯一因素。单纯为了迎合搜索引擎而选择与网站不相关的域名类型是一种得不偿失的做法。

7. 国家（地区）域名的选择

除了国际域名以外，几乎每个国家（地区）都会有各自的域名。例如中国大陆的国家域名包括.org、com.cn 等。

以 Google 为例，在世界上多个国家或者地区都有不同语言版本的 Google。对于来自不同国家（地区）的请求，它会优先返回与发出查询请求者所在地区最接近的网站。如使用英国版 Google（www.google.co.uk）查询 SEO，那么在搜索结果中，它会优先返回英国地区的网站。

从上面的例子中可以知道，搜索引擎除了可以根据网站的服务器 IP 及语言来识别网站所在的地区外，还会根据网站所使用的域名后缀来判断网站的地域性。所以，如果网站针对的是某个特定的国家（地区），那么在注册域名的时候应该优先使用这个国家（地区）的域名。但是，如果网站针对的是全球，那么就应该优先使用国际域名，如.com、.net 等。

做搜索引擎优化讲的是执行力，做的是“针线活”（做细节），对站内优化这一大环境来说，网站 URL 路径优化就是“针线活”，合理的 URL 路径有利于搜索引擎顺利抓取网页，反之会被搜索引擎淘汰，笔者认为对网站 URL 路径进行合理优化势在必行。

URL 路径分为绝对与相对路径、动态与静态路径。

绝对与相对路径。绝对路径简单来说体现的是真实路径，而相对路径体现的是路径之间的关系，由于相对路径在搜索引擎中表现良好，笔者建议尽量用相对路径。

动态与静态路径。搜索引擎的技术已经相对成熟，但笔者认为搜索引擎对静态路径的友

好度相对较高，建议使用静态路径，必要的时候可以 rewrite 路径，即伪静态。特殊情况下还可以用动态路径（因为有些空间不支持伪静态）。

URL 路径选择要点，路径尽量短，越短越好；针对动态路径，参数越少越好；路径要统一，网站中切勿出现动态路径和静态路径都被搜索引擎抓取的现象。

网站 URL 路径优化事项：

- 错误路径。尽量避免网站中出现不可以访问的 URL 路径，这样的路径对搜索引擎不友好。避免重复页面，这里的重复是搜索引擎中的体现，即相同文章内容不同的路径都被搜索引擎收录。
- URL 路径合理写法。笔者建议用全拼，拼音缩写，英文（这三种就要根据你是否做外贸，是否做品牌决定），URL 路径中字符建议用中划线“-”，尽量避免使用中文路径，笔者认为搜索引擎对中文路径还处于测试阶段。
- 目录与网站 URL 路径。必要的时候可以通过以把所有文章页面都放在一个目录里的方式实现路径的进一步合理优化，这样对站长来说方便管理，对搜索引擎来说方便抓取（路径短），对用户来说视觉效果好，就像百度百科那样把所有栏目中的大多数文章都放在目录 view 里面。
- URL 路径与网页收录。都知道中小型网站在三层以上的页面不会被搜索引擎收录，我们可以设置相应页面与权重高或者层级低的页面直接相连，如首页和栏目页，进而间接实现重要页面被抓取，最后得到很好的收录和排名。
- 切勿修改路径。在网站路径被搜索引擎收录以后千万不要修改路径，不然会产生死链接，但要变动的话，一定要将原来的 URL301 永远定向。

4.7 网页地址重定向

网页地址重定向是指当使用者浏览某个网址时，将他导向到另一个网址的技术。常常用在把一串很长的网址转成较短的网址上面。因为当要传播某网站的网址时，常常因为网址太长，不好记忆；又有可能因为换了网络的免费网页空间，网址又必须要变更，不知情的使用者还以为网站关闭了。这时就可以用网络上的转址服务了。常见的重定向有 301（永久重定向）及 302（暂时重定向）两种。

4.7.1 301 重定向

301 重定向是在搜索引擎优化中一个很常用的操作，可以帮助我们转移域名的权重。301 代表永久性转移（Permanently Moved），301 重定向是网页更改地址后对搜索引擎友好的最好方法，只要不是暂时转移的情况，都建议使用 301 来做转址。下面介绍几种常用的实

施方法。

1. Apache 服务器实现 301 重定向

在 Apache 中，有个很重要的文件.htaccess，通过对它的设置，可以实现很多强大的功能，301 重定向只是其中之一。

```
redirect permanent /index.php http://www.domain.com/index.php?
go=category_6（将网页 index.php 重定向到 http://www.domain.com/index.php?
go=category_6）
```

通过合理地配置重定向参数中的正则表达式，可以实现更复杂的匹配。

2. PHP 下的 301 重定向

```
<html>
<head>
<title>HTTP/1.1 301 Moved Permanently</title>
<?
Header ( "HTTP/1.1 301 Moved Permanently" ) ;
Header ( "Location: www.domain.com" ) ;
? >
</head>
<body>
HTTP/1.1 301 Moved Permanently
</body>
</html>
```

3. ASP 下的 301 重定向

```
<%@ Language=VBScript %>
<%
Response.Status="301 Moved Permanently"
Response.AddHeader "Location","http://www.domain.com"
%>
```

4. ASP .NET 下的 301 重定向

```
<script runat="server">
private void Page_Load(object sender, System.EventArgs e)
{
Response.Status = "301 Moved Permanently";
Response.AddHeader ("Location","http://www.domain.com");
}
</script>
```

5. ColdFusion 下的 301 重定向

```
<.cfheader statuscode="301" statustext="Moved permanently">
```

<.cfheader name="Location" value="http://www.domain.com">旧域名重定向到新域名创建一个.htaccess 文件，并将下面提供的代码写入文件内，它可以确保旧域名所有的目录或者网页正确的跳转到新域名内。

记住.htaccess 文件一定要放在旧网站的根目录下，并且新网站要和旧网站保持相同的目录结构及网页文件。

```
Options +FollowSymLinks
RewriteEngine on
RewriteRule (.*) http://www.domain.com/$1 [R=301,L]
```

请将上面的 www.domain.com 修改为你想要跳转到的域名。

此外，笔者建议大家归总旧网站的外部链接，并联系相应的站点修改导入链链的 URL，以指向新站点。

注意，.htaccess 文件只有在使用安装有 Mod Rewrite 模块的 Apache 作为 Web 服务器的 Linux 主机上才能起作用。

4.7.2 302 重定向

302 重定向又称之为 302 代表暂时性转移（Temporarily Moved ）302 redirect。 也被认为是暂时重定向（temporary redirect），一个暂时重定向是一种服务器端的重定向，能够被搜索引擎蜘蛛正确地处理。

4.8 页面静态化

为什么要把动态网页以静态网页的形式发布呢？其中有一个很重要的原因，就是因为搜索引擎。由于搜索引擎对 ASPX 页面收录和 HTML 页面收录率的差别以及页面资源占用问题，所以很多时候需要实现 ASPX 页面动态转静态。

以目前互联网上最普通的查找资料的方式看，互联网内容到达公众面前大致只有以下三条路。

- 第一条路是通过众所周知的门户网站，老读者有意识地到达这些地方搜索相应的有针对性的资料。
- 第二条路是做广告，通过公众使用一些免费服务的机会夹带资料信息，把公众带到目

标网站。

- 第三条路是搜索引擎，通过关键字和其他因素，把相应的信息推介到公众的面前。

实际上，即使是对门户网站，这也是带来新公众的最有效的途径；而对于绝大多数的站点，除了屈指可数的几个门户外，这种搜索引擎带来的读者公众，占了至少 90%以上的比例。如此重要的一条公众途径自然是不能忽略的，因此，尽管搜索机器人有点讨厌，各个网站不但不会再像从前一样把它封起来，反而热情无比地搞 SEO，所谓的面向搜索引擎的优化，其中就包括访问地址的改写，令动态网页看上去是静态网页，以便更多更大量地被搜索引擎收录，从而最大限度地提高自己的内容被目标接收的机会。

但是，在完全以动态技术开发的网站，转眼中要求变换成静态网页提供，同时，无论如何，动态网页的内容管理功能也是必须保留的；就如同一辆飞驶的奔驰忽然要求 180°转弯，要付出的成本代价是非常大的，是否真的值得，也确实让人怀疑。

另一个重要原因就是提高程序性能。很多大型网站，进去的时候看它很复杂的页面，但是加载也没有耗费多长时间，除了其他必要原因以外，静态化也是其中必须考虑的技术之一。它先于用户获取资源或数据库数据进而通过静态化处理，生成静态页面，所有人都访问这一个静态页面，而静态化处理的页面本身的访问速度要较动态页面快很多倍，因此程序性能会有大大的提升。总之静态化在页面上的体现为：访问速度加快，用户体验性明显提升；在后台体现为：访问脱离数据库，减轻了数据库访问压力。

静态化虽然可以提升程序的性能，但是它并不是提升整体性能的根本原因，就像一台电脑，只有 CPU 好、显卡好，内存好，是不行的，要看整体性能。很多时候是开发者的原因，导致程序自身性能不佳，所以性能量力而行，以项目本身性能为基础，其他优化方法为辅，最终提升整个应用程序的性能。

4.9 小结

网站 URL 路径优化就是站内优化的冰山一角，做 URL 优化就是做“针线活”，只有缝好每一针才能体现其真正的意义，总之经过合理优化后的网站 URL 路径既满足用户体验，又对搜索引擎好处。

第 5 章　页面代码优化

页面是网站建设中的基本组成部分，代码优化则是网站页面优化中的一个重要环节，也是页面优化的基础，SEO 搜索引擎优化要做的基本工作就是改善网页代码。本章将具体介绍页面的代码优化，主要包括页面代码的精简方法，HTML 标签优化，标头优化，权重标签的使用，图片的优化以及沙盒效应知识。

5.1 页面代码精简

代码精简就是指清除和简化网站页面中的代码，使之降低页面的体积、提高网站的用户体验。如果我们对网站代码进行精简和优化，以及对网页代码中的标签进行适当的优化，有助于蜘蛛的爬行和抓取，也有助于增加空间访问速度。这当中就要涉及到 CMS。

1. 什么是 CMS 程序

CMS 是 Content Management System 的缩写，意为“内容管理系统”，它具有许多基于模板的优秀设计，可以加快网站开发速度和减少开发成本，是现在网络上普遍运用的快速建立网站的方法。

CMS 的设计出发点其实就是方便一些对网络编程、脚本代码并不是很熟悉的用户，用一种比较简单的方式来建立、管理、运营自己的网站。CMS 的功能并不只限于文本处理，它也可以处理图片、Flash 动画、声像流、图像甚至电子邮件档案。

在互联网的初期，我们看到的网站基本都是采用 HTML 编码方式，采取手动更新页面、手工填充信息，非常麻烦。后来逐渐过渡到使用 ASP、PHP、JSP 等动态脚本编写网站的阶段，这时的网站通过脚本和数据库的交互，实现网站数据快速、方便地发布和管理。但这时网站脚本程序的开发往往需要投入很大的成本，一般只有大型站点才有自己开发的网站脚本程序。现在，随着 CMS 程序的出现、发展和开源，任何网民如果想要建立自己的网站，都可以通过使用 CMS 的方式，快速、简便地建立一个从功能上完全可以比拟门户网站的站点。

简而言之，CMS 就是可以让网站建设者不需要学习繁杂的建站技术，不需要学习太多复杂的脚本语言，就能够构建出一个风格统一、功能强大的专业网站平台。

2. CMS 分类

很多新手站长通常会把 CMS 分类弄混，这是因为 CMS 分类方式有很多种，下面简单介绍最常见的应用类别和脚本类型的分类方法。

根据应用层面的不同，CMS 可以被划分为：重视后台管理的 CMS、重视风格设计的 CMS 和重视前台发布的 CMS。

根据脚本语言的种类，CMS 又可以划分为：ASP 类的 CMS、PHP 类的 CMS 和 JSP 类的 CMS。

通常在搜索引擎优化者眼中，为了方便使用和交流，CMS 基本是按照脚本代码的类型来进行划分的。

就目前已经存在的各种 CMS 来说，界面都大同小异，但是功能却有不小的差异，在编程风格与管理方式上也是千差万别的。另外，由于各个 CMS 系统的原创者自己本身的背景不同，对用户体验的理解不同，对 SEO 的功能需求支持不同，所以现在的 CMS 仍然没有统一的标准。

5.1.1 网页代码对蜘蛛的影响

我们先说明一个关于代码优化的误区问题然后再来探讨网页代码优化，略懂 SEO 的个人站长或优化团队会推荐页面采用完全 DIV+CSS 样式来提高网页速度，说：“对搜索引擎收录友好、现在很流行”，甚至千方百计避免在网页中使用 TABLE。实际上 DIV+CSS 只是 HTML 网页布局方法中的其中一种，最大的优点在于用户体验比较好，因为 DIV+CSS 可以在网页 HTML 没有完全下载完毕的情况下就能展现一部分已下载完成的网页 DIV 区域内容。这不是说搜索引擎也获取 DIV+CSS 页面速度比较快，因为搜索引擎需要的是等该网页完全下载完毕储存在数据库才能开始分析，不会先分析一部分等会再分析，用户体验的网页打开呈现速度快不等于整个网页下载速度快，网页下载速度快只有通过优化网页代码、内容来解决。

一般情况下，TABLE 需要把 TABLE 下载完毕才能显示，当然也可以不考虑外观的情况下用 CSS 样式来解决 TABLE 不完全下载就能显示的问题。所以 DIV+CSS 是不能改善“搜索引擎体验”的，DIV+CSS 流行的也不见得是有用的。这和一些网页甚至用户都浏览不了但会被搜索引擎收录的原因是一样的。如果说要针对用户体验优化（用户体验优化包含网页 SEO），SEO 优化建议网页在相对不增加网页代码的情况下，更多采用 DIV+CSS 布局来设计。

DIV+CSS 是一种网页的布局方法，这一种网页布局方法有别于传统的 TABLE 布局，DIV+CSS 和 TABLE 都符合 W3C 标准，只要符合 W3C 标准都可以保证你的网站不会因为将来网络应用的升级而被淘汰。DIV+CSS 对搜索引擎优化与否取决于网页设计的专业水平，而不是 DIV+CSS 本身；很多 DIV+CSS 设计的网页并不能保证网页对搜索引擎的优化，甚至不能保证一定比 TABLE 网页有更简洁的代码设计，如 4 列或 4 列以上网页结构的布局看看到底是 DIV+CSS 简单还是 TABLE 简单，多列的网页布局结构会使 DIV+CSS 设计变得异常困难，DIV+CSS 布局一般只用在相对简单网页上。

搜索引擎对于网页的收录和排序，不是以是否采用 TABLE 和 DIV+CSS 定位来衡量的，也重来没有一个搜索引擎表示会对 DIV+CSS 友好，说只有 DIV+CSS 布局对搜索引擎友好的只是一部分不负责任的伪 SEOer 意想罢了。通过观察搜索引擎并没有对网页采用何种布局的收录或排名有影响，倒是网页内容、关键字在网页中的分布、相关网站链接对搜索引擎的排序影响比较大。

DIV+CSS 和 TABLE 在网页设计中各有优缺点，SEO 网页代码优化应灵活运用使网页结构和代码简单、关键字在网页内容中分布均匀、网页开发和维护简单。

DIV+CSS 有以下优点：

- 网页载入比较快，由于 DIV+CSS 中的 CSS 富含丰富的样式，使网页定位和表现更加灵活。
- 页面内容与样式分离可以使网页代码减少，使页面和样式的调整变得更加方便。
- 表现和结构分离，在团队开发中更容易分工合作而减少相互关联性。

DIV+CSS 缺点：

- DIV+CSS 到现在还没有实现所有浏览器的统一兼容，使得设计变得更复杂。
- DIV+CSS 比 TABLE 定位复杂得多，多层嵌套的 DIV 会严重影响网页代码的可阅读性，对于网页设计高手也很容易出现问题，更不要说初学者了，CSS 通常放在一个或几个外部文件中，有可能相当复杂和比较庞大，如果 CSS 文件调用异常那么整个网页将变得惨不忍睹。

TABLE 有以下优点：

- TABLE 表格对浏览器兼容性非常好，是传统的 HTML 元素，在可以预计的将来，TABLE 的地位依然十分重要。
- 对网页内数据的呈现更简单、对于动态网页的设计开发更加方便、网页代码更简洁。
- TABLE 在网页中的定位相对 DIV+CSS 比较方便。

TABLE 缺点：

- 复杂的网页结构中 TABLE 对于网页设计不够灵活。
- TABLE 的嵌套会大大增加网页代码，影响网页代码的可阅读性。
- 网页中较多的 TABLE 和 TABLE 中较多的数据内容会影响网页呈现速度（只有 TABLE 加载完毕才能呈现）。
- 网页更新和维护比较复杂、网页的样式调整比较困难（这也是通用网站程序不采用 TABLE 的主要原因之一）。

SEO 网页代码优化后使网页中主要内容在设计时要注意放在网页代码的上方，搜索蜘蛛一般是获取网页内代码上方并且连续的文字会作为该页的主要内容。但 DIV+CSS 布局往往不会满足这种设计要求，在这种情况下，可以使用 TABLE 或在网页代码上部增加该页内容

的概括文字来均匀分布网页关键字。

5.1.2 DIV+CSS 的优势

首先要符合 W3C 标准。现在微软等公司均为 W3C 支持者。这一点是至关重要的，因为这可以保证你的网站不会因为将来网络应用的升级而遭到淘汰。支持浏览器的向后兼容，也就是无论在未来的浏览器大战中，胜利的是 IE 7 还是火狐，你的网站都能很好的兼容，搜索引擎也会更加友好。相对与传统的 TABLE，采用 DIV+CSS 技术的网页，对于搜索引擎的收录更加友好。

样式的调整更加方便。内容和样式的分离，使页面和样式的调整变得更加方便。 现在 Yahoo！、MSN 等国际门户网站，网易、新浪等国内门户网站，和主流的 Web 2.0 网站，均采用 DIV+CSS 的框架模式，更加印证了 DIV+CSS 是大势所趋。

CSS 的极大优势表现在简洁的代码。对于一个大型网站来说，可以节省大量带宽，而且众所周知，搜索引擎喜欢简洁的代码。表现和结构分离，在团队开发中更容易分工合作而减少相互关联性。

5.1.3 CSS 放入网页的方式

行内套用是指在网页其中的一行套入其他的行或者对象，我们可以在 HTML 文件内直接宣告样式。

例如 1：

```
Thisisfontsize16
```

以上的 HTML 代码在浏览器上会显现为：Thisisfontsize16。

嵌入套用样式可以嵌入于 HTML 文件中（通常是在<head>内）。

例如 2：

```
<head><styletype="text/css">div{div{background-color:#
FF0000;}</style></head><body>背景颜色是红色</body>
```

以上的 HTML 会显现出：背景颜色是红色。外部连接套用在这种方式下，所有的 CSS 样式宣告都是存在另外一个档案中。

该档案通常名称为.css。

在 HTML 文件的<header>…</header>之中，将用以下的程式码将这个.css 档案连接进入：

```
<linkrel=stylesheettype="text/css" href="external-stylesheet.css">
```

以上这一行会将在 external-stylesheet.css 档案内所宣告的样式加入 HTML 文件内。汇入

套用外部的 CSS 样式也可以被汇入进 HTML 文件。

汇入的做法为利用@import 指令。

@import 的语法为：

```
<STYLETYPE="text/css"><STYLETYPE="text/css"><</STYLE>@imp ort">!--
@importurl ( http://www.divhome.com/#.mysite.com/style.css) ;@im porturl
(</STYLE>
```

@import 指令最初的用意，是为了能够针对不同的浏览器而运用不同的样式。不过，现在已经没有这个必要。现在用@import 的目的，是要加入多个 CSS 样式。当多个 CSS 样式被@import 的方式加入，当不同的 CSS 样式互相有冲突时，后被加入的 CSS 样式有优先的顺位（详情请见 CSS 串接）。

5.1.4 DIV+CSS 最常用的布局标签

估计使用<h1>标签的人寥寥无几，因为它显示的字体超级大，但我们是 CSSer，有什么标签样式不能改呢？而且它能表述的意义就是它本身的作用，该标签在大多数情况下都是用来定义导航区域的，当然也可以用<ol>标签来替代，因为导航连接没有什么顺序，所以还是用<ul>标签比较准确（去掉 CSS 后效果更显著）。

<b>这个已经不再推荐使用的标签，在布局上却能带来不少的方便（因为短），用在细小地方的布局定义还是不错的选择。

<h2>使用最多的地方应该是副标题，但有些地方需要定义栏目样式的话，用这个标签比较合适，栏目内容就使用<p>标签。

5.1.5 使用 DIV + CSS 布局网站的优点

随着 Web 2.0 标准化设计理念的普遍，国内许多大型门户网站已经纷纷采用 DIV+CSS 的制作方式，网络也成为 CSS 网页布局的技术学习和先锋站点，以实际应用状况来看，此种方法绝对大于表格制作页面的方法。

如今大部分网站仍然采用表格嵌套内容的方式来制作网站，虽然此方法对于我们来说比较熟悉、比较容易上手；但是，它却阻碍了一种更好的、更有亲和力的、更灵活的，而且功能更强大的网站设计方法：DIV+CSS。

5.1.6 CSS 网页布局的意义体现

（1）可以使页面载入得更快。由于将大部分的页面代码写在了 CSS 当中，所以使得页面体积容量变得更小。相对于表格嵌套的方式，DIV+CSS 将页面独立成更多的区域，在打开页面的时候，逐层加载。而不像表格嵌套那样将整个页面圈在一个大表格里，使得加载速

度很慢。

（2）降低流量费用，页面体积变小，浏览速度变快，这就使得对于某些控制主机流量的网站来说是最大的优势了。

（3）修改设计时更有效率。由于使用了 DIV+CSS 制作方法，在修改页面的时候就更加省时。根据区域内容标记，到 CSS 里找到相应的 ID，使得修改页面的时候更加方便，也不会破坏页面其他部分的布局样式。

（4）保持视觉的一致性。以往表格嵌套的制作方法，会使得页面与页面，或者区域与区域之间的显示效果会有偏差。而使用 DIV+CSS 的制作方法，将所有页面，或所有区域统一用 CSS 文件控制，就避免了不同区域或不同页面体现出的效果偏差。

（5）更好地被搜索引擎收录，由于将大部分的 HTML 代码和内容样式写入了 CSS 文件中，这就使得网页中正文部分更为突出明显，便于被搜索引擎采集收录。

（6）对访问者和浏览器更具亲和力。我们都知道网站做出来是给访问者使用的，DIV+CSS 在这方面更具优势。由于 CSS 富含丰富的样式，使页面更加灵活性，它可以根据不同的浏览器，而达到显示效果的统一和不变形。

5.1.7 CSS 网页布局的副作用

说了这么多 CSS 网页布局的意义与优点，同时也不能轻视 CSS 网页布局的副作用。

- DIV+CSS 尽管不是高不可及，但至少要比表格定位复杂得多，即使对于网站设计高手也很容易出现问题，更不要说初学者了。
- CSS 网站制作的设计元素通常放在一个外部文件或几个文件中，有可能相当复杂，甚至比较庞大，如果 CSS 文件调用出现异常，那么整个网站将变得惨不忍睹。
- 虽然说 DIV+CSS 解决了大部分浏览器兼容问题，但是也有在部分浏览器中使用出现异常，比如火狐浏览器，在 IE 中显示正常的页面，到了火狐浏览器中可能会面目全非。当然这应该是浏览器的问题，但是可以说在目前来看，DIV+CSS 还没有实现所有浏览器的统一兼容。
- DIV+CSS 对搜索引擎优化与否，取决于网页设计的专业水平，而不是 DIV+CSS 本身。DIV+CSS 网页设计并不能保证网页对搜索引擎的优化，甚至不能保证一定比 HTML 网站有更简洁的代码设计，何况搜索引擎对于网页的收录和排序显然不是以是否采用表格和 CSS 定位来衡量，这就是为什么很多传统表格布局制作的网站在搜索结果中的排序靠前，而很多使用 CSS 及 Web 标准制作的网页排名依然靠后的原因。因为对于搜索引擎而言，网站结构、内容、相关网站链接等因素始终是网站优化最重要的指标。

笔者认为不可盲目跟风，如何更有效、更合理的运用 Web 2.0 设计标准，这需要很长时间的学习和锻炼。而如何将 DIV+CSS 运用得更好，笔者觉得这需要通过不断的实践和体检，积累丰富的设计经验，才能很好的掌握这门技术。

5.1.8　DIV+CSS 常见错误总结

CSS＋DIV 是网站标准（或称“Web 标准”）中常用的术语之一，通常为了说明与 HTML 网页设计语言中的表格（TABLE）定位方式的区别，因为 XHTML 网站设计标准中，不再使用表格定位技术，而是采用 DIV+CSS 的方式实现各种定位。应用 DIV+CSS 编码时很容易犯一些错误。这里列举了一些常见的错误。

（1）检查 HTML 元素是否有拼写错误、是否忘记结束标记

即使是老手也经常会弄错 DIV 的嵌套关系。可以用 Dreamweaver 的验证功能检查有无错误。

（2）检查 CSS 是否正确

要检查有无拼写错误、是否忘记结尾的“}”符号等，可以利用 CleanCSS 来检查 CSS 的拼写错误。因为 CleanCSS 本是为 CSS 精简压缩的工具，但也能检查出拼写错误。

（3）确定错误发生的位置

如果错误影响了整体布局，则可以逐个删除 DIV 块，直到删除某个 DIV 块后显示恢复正常，即可确定错误发生的位置。

（4）利用 border 属性确定出错元素的布局特性

使用 float 属性布局一不小心就会出错。这时为元素添加 border 属性确定元素边界，错误原因即可水落石出。

（5）float 元素的父元素不能指定 clear 属性

在 MacIE 下如果对 float 的元素的父元素使用 clear 属性，周围的 float 元素布局就会混乱。这是 MacIE 的著名的 bug，倘若不知道就会走弯路。

（6）在 float 元素中务必指定 width 属性

很多浏览器在显示未指定 width 的 float 元素时会有 bug。所以不管 float 元素的内容如何，一定要为其指定 width 属性。另外指定元素时尽量使用 em 而不是 px 做单位。

（7）float 元素不能指定 margin 和 padding 等属性

IE 在显示指定了 margin 和 padding 的 float 元素时有 bug。因此不要对 float 元素指定 margin 和 padding 属性（可以在 float 元素内部嵌套一个 DIV 来设置 margin 和 padding）。也可以使用 hack 方法为 IE 指定特别的值。

（8）float 元素的宽度之和要小于 100%

如果 float 元素的宽度之和正好是 100%，某些古老的浏览器将不能正常显示。因此请保证宽度之和要小于 99%。

（9）是否重设了默认的样式

对某些属性如 margin、padding 等，不同浏览器会有不同的解释。因此最好在开发前首先将全体的 margin、padding 设置为 0、列表样式设置为 none 等。

（10）是否忘记了写 DTD

如果无论怎样调整不同浏览器，其显示结果还是不一样，则可以检查页面开头是否忘了写 DTD：<!DOCTYPE HTML PUBLIC “-//W3C//DTD HTML 4.01 Transitional//EN”。

5.2 HTML 标签优化

既然要削减代码体积，那就先从没必要保留的地方下手，也许你是用的智能编辑器如 Dreamweaver、VS 等，让你点点鼠标就能插入代码，但这些代码中也有一些代码是不必要的。下面是一些常见的可优化的 HTML 标签。

1. 去除不必要的 HTML 标签

<meta>标签。除了其中的编码声明行、<keyword>与<description>子标签，其余的都可以删除。

<link>标签。通常用来载入 CSS 文件，其余用途则可删除；而且，link 载入 CSS 时，带上 type=“text/css”即可，不用再加上 rel=“stylesheet”，media 属性也可不要。

通常<html xmlns=“http://www.w3.org/1999/xhtml”>这句中只保留<html>即可。

<script>标签。只须保留 type，而无须使用 language 属性。

内容为空的属性。比如 img 的空白 alt，a 的空白 title。

加在不需要视觉操作的元素上的属性。如 link，cript 上的 class，id 等。

2. 主动减少 HTML 标签

这点需要有比较良好的代码知识。我们应该本着写更少的代码，却能做同样多事情的心态来写 HTML。

重点：减少 HTML 标签嵌套。就是能用一层标签的就不要用两层，然后通过 CSS 来达到相同的视觉效果。

该块的 HTML 结构和写法如下：

```
<div>
<div><h2>相关文章</h2></div>
<div><ul><li>。。。</li></ul></div>
</div>
```

也可以是如下写法：

```
<div>
<h2>相关文章</h2>
<ul><li>。。。</li></ul>
</div>
```

相比之下，后者代码精简得多，而且可以通过 CSS 实现相同的视觉效果。减少嵌套，对 SEO 也更好。

不过，前者无疑具有更好的扩展性。如你想在“相关文章”右侧加一个“更多”链接，那上述这种结构就很容易修改。

所以，精简代码，是在确认了需求改动不频繁的前提下，否则以后修改起来会让你很痛苦。

3. 尽量少使用行内 CSS

因为行内 CSS 每次都要加个 style=，关键是，只能对此元素生效。不如用一个 class 代替，重用起来也方便。把不需要判断条件载入的 CSS 文件给合并了，可以减少一个<link>标签，也就减少了一个 http 请求。

5.2.1　垃圾代码清理

清理垃圾代码是指删除页面中的多余代码。一个页面经过垃圾代码清理后，可以删除接近 80%的多余代码。然而，什么样的代码才是垃圾代码呢？

1. 垃圾代码

垃圾代码就是指删除后不会影响页面正常显示的非必要的代码。这些代码占据着庞大的空间，不但消耗用户下载页面的时间，还占用搜索引擎巨大的存储空间，增加搜索引擎分析数据的时间。然而，这么庞大的垃圾代码是如何产生的呢？

2. 垃圾代码的产生

一般情况下，我们会使用网页制作软件，如 FrontPage、Word 及 Dreamweaver。这些网页制作软件在制作网页的时候，会生成相应的 HTML 代码，其中大部分是必要的，而有一小部分则是可有可无的，这就是所谓的垃圾代码。因此，避免产生垃圾代码最好的办法就是以手工编写代码的方式制作网页。

即使是同样内容的网页，使用不同的网页制作软件进行制作，所产生的代码也会存在很大的区别。在没有掌握手工编写代码前，可以使用 Dreamweaver 制作网页，完成后再清理其

中的垃圾代码。

3. 垃圾代码的清理

空格字符是网页中最常见的垃圾代码。但此处讨论的空格字符并非 HTML 语言中的 < > 标签，而是指在代码编辑环境下敲击键盘上的空格键所产生的符号。网页中每个空格都相当于一个字符，因此，空格字符也会占用巨大的空间。初步统计，空格字符约占用页面总体积的 1.5%。也就是说，一个 100KB 的页面中，其中 1.5KB 是空格字符。

空格字符通常会出现在每行代码的开始和结束处，还有就是空白行。

- 每行代码开始前的空格字符。是网页制作软件生成代码时，默认产生的缩进字符。
- 每行代码结束处的空格字符。是编辑 HTML 代码时，不小心加上的。只要我们在每行代码的结束位置向右移动光标，如果光标不换行，则说明该页面还存在空格字符。
- 空白行。在编写代码时，常利用空行或者注释对不同功能模块进行分隔，这样就产生了空白行。

对于页面中不同位置上的空格字符，清理的方式也有所不同，下面介绍几种常用的空格字符清理方法。

（1）每行代码开始前的空格字符

对于每行代码开始前的空格字符，可以使用 Dreamweaver 或者 EditPlus 等软件进行清理。具体操作步骤如下：

步骤 01 转至 HTML 代码编辑模式（如果是 Dreamweaver）。

步骤 02 按 Ctrl+A 组合键选择全部代码。

步骤 03 使用 Shift+Tab 组合键，直至删除所有空格为止。

下面的代码 1 是清理空格前的效果，而代码 2 则是清理空格字符后的效果。

代码 1：

```
<table width="100%" border="0" cellsoacing="0" cellpadding="0">
<tr>
<td> </td>
</tr>
<tr>
<td> </td>
</tr>
<tr>
<td> </td>
</tr>
</table>
```

代码 2：

```
<table width="100%" border="0" cellsoacing="0" cellpadding="0">
<tr>
<td> </td>
</tr>
<tr>
<td> </td>
</tr>
<tr>
<td> </td>
</tr>
</table>
```

（2）空白行

对于空白行中空格字符的清理，只能手工逐行去掉。

步骤 01 首先，使用鼠标选择空行。

步骤 02 再按 Delete 键，即可删除空白行。

（3）每行代码结尾处的空格字符

对于每行代码结尾处的空格字符，可以使用 Editplus 或者其他文笔编辑软件的替换功能进行清理。

步骤 01 用鼠标选择“>”+“一个空格字符”，即“>”。

步骤 02 再按 Ctrl+F 组合键，就会弹出一个窗口。

步骤 03 在“替换”字段输入“>”，再单击“全部替换”即可。

页面经过清理空格字符后，大概可以删除 70%左右的垃圾代码。也就是说，空格字符的代码量大概占总垃圾代码的 70%，还有 30%左右的垃圾代码分布在默认属性、注释语句及空语句中。

（1）默认属性

在利用网页制作软件制作网页时，通常会产生一些默认属性的代码。如在 HTML 代码中不添加左对齐属性，页面中的内容也是以左对齐的方式显示的，所以代码中的左对齐属性是可以删除的。

在页面中，还有很多其他属性都是默认属性，以下是比较常见的。

- align=“left”：横向居左对齐属性值，默认情况下文章或者图片都是横向居左对齐的。
- valign=“middle”：竖向居中对齐属性值，默认情况下文章或者图片都是竖向居中对齐的。

- size= “3” ：文字大小属性值，默认情况下是 3 号字体。
- target= “_self” ：新页面打开属性值，默认情况下网页中文本的字体颜色是黑色。

利用 Dreamweaver 的替换功能，可以完成对默认属性代码的替换。在“查找”中填上要替换的默认属性代码，“替换”处留空白，还可以在“查找范围”上选择替换范围，然后点击“替换全部”就可以删除“产照范围”内所有指定的默认属性代码。

（2）注释语句

注释语句是用于对代码功能或作用进行说明的语句，注释语句里的内容对于普通用户来说就是不可见的。即普通用户在页面中是看不到 HTML 代码中注释标签里的内容。

在制作网页的时候，都会添加注释语句增强代码的可读性。但是，过多的注释语句会占用大量的存储空间。不仅如此，如果在注释标签里添加大量的关键字，还会被搜索引擎认为是堆砌关键字，从而对网站进行惩罚。

所以，在添加注释语句时，绝不添加那些即使删除后也不会影响网页源代码可读性的注释语句。

（3）空语句

垃圾代码清理的最后一步就是删除代码中的空语句。简单地说，空语句就是指标签间不存在任何内容的语句。

空语句的清理主要针对即使删除后也不会影响页面正常显示的标签。但有一部分标签是不能删除的（如<tr>、<td>），否则页面将会出现错乱。

手工编写 HTML 代码是搜索引擎优化从业者必须具备的基本技能。如果你已经具备手工编写 HTML 代码的能力，就应该有限采用手工编写代码的方式制作网页。

5.2.2 JavaScript 优化

JavaScript 是一门解释性的语言、它不像 java、C#等程序设计语言由编译器先进行编译再运行，而是直接下载到用户的客户端进行执行。因此代码本身的优劣就直接决定了代码下载的速度以及执行的效率。

1. 减缓代码下载时间

Web 浏览器下载的是 JavaScript 的源代码，其中包含的长变量名、注释、空格和换行等多余字符大大减缓了代码下载的时间。这些字符对于团队编写是十分有效，但在最后工程完成上传到服务器时，应当将它们全部删除。

例如：

```
function showMeTheMoney () { if (!money) { return false; }else{ ... } }
```

可优化成：

```
function showMeTheMoney () {if (!money) {return false;}else{...}}
```

这样优化后就节约了 25 个字节、倘若是一个大的 JavaScript 工程、将节省出非常大的空间、不但提高了用户的下载速度、也减轻了服务器的压力。

另外，对于布尔型的值 true 和 false，true 可以用 1 来代替，而 false 可以用 0 来代替。对于 true 节省了 3 个字节，而 false 则节省了 4 个字节。

例如 1：

```
var bSearch = false;
for (var i=0;i<aChoices.length&&!bSearch;i++) { if (aChoices[i] == vValue)
bSearch = true ; }
```

替换成：

```
var bSearch = 0;
for (var i=0;i<aChoices.length&&!bSearch;i++) { if (aChoices[i] == vValue)
bSearch = 1 ; }
```

替换了布尔值之后，代码的执行效率和结果都相同，但节省了 7 个字节。

代码中常常会出现检测某个值是否为有效值的语句，而很多条件一定要判断某个变量是否为“undefined”、“null”、或者“false”。

例如 2：

```
if (myValue != undefined) { //... } if (myValue !=null) { //... } if
(myValue != false) { //... }
```

这些虽然都正确、但采用逻辑非操作符"!"也可以有同样的效果。

例如 3：

```
if (!myValue) { //... }
```

这样的替换也可以节省一部分字节，而且不影响代码的可读性。这类型的代码优化还有将数组定义时的 new Array () 直接用“[]”代替、对象定义时的 new Object () 用“{}”代替等。

例如 4：

```
var myArray = new Array () ;
var myArray = [];
var myObject = new Object () ;
var myObject = {};
```

显然，第二行和第四行的代码较为精简，而且也很容易理解。

另外，在编写代码时往往为了提高可读性，将函数名称和变量名称都使用了很长的英文单词，同时也大大增加了代码的长度。

例如5：

```
function
AddThreeVarsTogether (firstVar,secondVar,thirdVar) { return
(firstVar+secondVar+thirdVar) ; }
```

可优化成：

```
function A (a,b,c) {return  (a+b+c) ;}
```

注意，在进行变量名称替换时、必须十分小心、尤其不推荐使用文本编辑器的“查找”、“替换”功能，因为编辑器不能很好地区分变量名称或者其他代码。

例如6：

希望将变量“tion”全部替换成“io”，很可能导致关键字“function”也被破坏。

对于上述这些减少代码体积的方法，有一些很实用的小工具可以自动完成类似的工作。如：ECMAScript Cruncher、JSMin、Online JavaScript Compressor等。

2. 合理声明变量

减少代码的体积仅仅只能使得用户下载的速度变快、但执行程序的速度并没有改变。要提高代码执行的效果、还得在各方面做调整。

在浏览器中、JavaScript默认的变量范围是window对象、也就是全局变量。全局变量只有在浏览器关闭时才释放。而JavaScript也有局部变量、通常在function中执行完毕就会立即被释放。因此在函数体中要尽可能使用var关键字来声明变量：

```
function First () { a = "" ; //直接使用变量 } function Second () { alert
(a) ; } First () ; Second () ;
```

这样，变量“a”就成为了全局变量、直到页面关闭时才会被销毁、浪费了不必要的资源，如果在“a”的前面加上“var”，这样“a”就成为了当前function的局部变量。在执行完First（）便立即被销毁。因此，在函数体中如果不是特别需要的全局变量，都应当使用“var”进行声明，从而节省系统资源。

3. 使用内置函数缩短编译时间

只要可能就应当尽量使用JavaScript内置的函数。因为这些内置的属性、方法都是用C或者C++语言编译过的，运行起来比实时编译的JavaScript快很多。如计算指数函数，可以自己编写：

```
<html>
<head>
```

```
<base href="<%=basePath%>">
<title>内置函数</title>
<meta http-equiv="pragma" content="no-cache">
<meta http-equiv="cache-control" content="no-cache">
<meta http-equiv="expires" content="0">
<meta http-equiv="keywords" content="keyword1,keyword2,keyword3">
<meta http-equiv="description" content="This is my page">
<script type="text/javascript">
function myPower (iNum,n) {
var iResult = iNum ;
for (var i=0;i<n;i++)
iResult *= iNum ;
return iResult;
}
var myDate1 = new Date () ;
for (var i=0;i<150000;i++) {
myPower (7,8) ; //自定义方法 }
var myDate2 = new Date () ;
document.write (myDate2 - myDate1) ;
document.write ("<br/>") ;
myDate1 = new Date () ;
for (var i=0;i<150000;i++) {
Math.pow (7,8) ; //采用系统内置方法
}
myDate2 = new Date () ;
document.write (myDate2 - myDate1) ;
</script>
</head>
<body>
</body>
</html>
```

运行的结果是：自定义方法用了 15，内置方法用了 1（不同的计算机运行速度会有差别）这样就能看出，系统内置的方法要快很多。

4. 合理书写 if 语句

if 语句恐怕是所有代码中使用最频繁的，但是它的执行效率并不是很高。在用 if 语句和多个 else 语句时，一定要把最有可能的情况放在第一个，然后是可能性第二的，依此类推。如预计某个数值在 0~100 之间出现的概率最大，则可以这样安排代码：

```
if (iNum>0&&iNum <100) {
alert ("在 0 和 100 之间") ;
}
else if (iNum>99&&iNum<200)
{
```

```
alert ("在 100 和 200 之间") ;
}
else if (iNum>199&&iNum<300) {
alert ("在 200 和 300 之间") ;
}else
{
alert ("小于等于 0 或者大于等于 300") ;
}
```

总是将出现概率最多的情况放在前面，这样就减少了进行多次测试后才能遇到正确条件的情况。当然也要尽可能减少使用 else if 语句，例如：以上的代码还可以进一步优化为如下代码：

```
if (iNum>0) {
if (iNum<100) {
alert ("在 0 和 100 之间") ;
}
else{
if (iNum<200) {
alert ("在 100 和 200 之间") ;
}
else{ if (iNum<300) {
alert ("在 200 和 300 之间") ;
}else{
alert ("大于等于 300") ;
} } } }
else{ alert ("小于等于 0") ;
}
```

以上的代码看起来比较复杂，但因为考虑了很多代码潜在的判断问题、执行问题，因此执行速度要较前面的代码快。

另外，通常当超过两种情况时，最好能够使用 switch 语句。经常用 switch 语句代替 if 语句，可令执行速度快 10 倍左右。另外，由于 case 语句可以使用任何类型，也大大方便 switch 语句的编写。

5. 最小化语句数量

脚本中的语句越少，执行的时间就越短，而且代码的体积也会相应减少。例如使用 var 定义变量时可以一次定义多个。代码如下：

```
var iNum = 365;
var sColor = "yellow";
var aMyNum = [8,7,12,3] ;
var oMyDate = new Date () ;
```

上面的多个定义可以用 var 关键字一次性定义。代码如下：

```
var iNum = 365, sColor = "yellow" , aMyNum = [8,7,12,3],oMyDate = new Date
()  ;
```

同样在很多迭代运算的时候，也应该尽可能减少代码量。如下两行代码：

```
var sCar = aCars[i]; i++;
```

可优化成：

```
var sCar = aCars[i++];
```

6. 节约使用 DOM

JavaScript 对 DOM 的处理可能是最耗费时间的操作之一。每次 JavaScript 对 DOM 的操作都会改变页面的表现，并重新渲染整个页面，从而有明显的时间消耗。比较快捷的方法就是尽可能不在页面进行 DOM 操作，如下例中为 ul 添加了 10 个条目。

```
var oUl = document.getElementById("ulItem");
for(var i=0;i<10;i++){
var oLi = document.createElement("li");
oUl.appendChild(oLi);
oLi.appendChild(document.createTextNode("Item "+i));
}
```

以上代码在循环中调用 oUl.appendChild（oLi），每次执行这条语句后，浏览器就会重新渲染页面，其次给列表添加文本节点 oLi.appendChild（document.createTextNode（"Item "+i）），这也会造成页面被重新渲染。因此每次运行都会造成两次重新渲染页面，共 20 次。

通常应当尽可能减少 DOM 的操作，将列表项目在添加文本节点之后再添加，并合理地使用 createDocumentFragment（）。代码如下：

```
var oUl = document.getElementById("ulItem");
var oTemp = document.createDocumentFragment();
for(var i=0;i<10;i++){ var oLi = document.createElement("li");
oLi.appendChild(document.createTextNode("Item "+i));
oTemp.appendChild(oLi);
}
oUl.appendChild(oTemp);
```

5.2.3　表格优化

精简代码的最后一步就是对表格进行优化。表格优化主要是解决表格之间的嵌套问题，让表格之间保持独立，从而加快页面在浏览器中的显示速度，提高用户体验。同时，又可以

精简页面代码，提高页面的搜索引擎友好性。

在介绍表格优化之前，我们先来了解浏览器解析 HTML 代码的过程。当浏览器的编译器遇到一个开始标签时，就会寻找与之匹配的结束标签，直至匹配才会在浏览器中显示该标签中的内容。如当浏览器的编译器遇到<table>标签时，就寻找与之匹配的结束标签</table>，匹配上才显示表格里的内容。否则将继续往下寻找，直至匹配为止。所以，如果表格间出现多次嵌套，特别是当这些表格里的内容较多的时候，打开页面的速度就会特别慢，因为需要下载及编译的内容特别庞大。

看下面这个例子。一个 3 行 1 列的大表格，每个单元格里嵌套着一个小表格。假设每个小表格内容的代码量有 50KB，则大表格里代码量至少有 150KB。也就是说，浏览器至少需要下载并执行 150KB 的代码后才能显示表格里的内容。换句话说，只有下载并执行 150KB 的代码后，这个页面才会显示。在目前的网速条件下，对于这么巨大的代码量，访问者显然是接受不了的。代码内容如下所示：

```
<table width="778" height="105" border="1" cellpadding="1"
cellspacing="1">
<tr>
<td>
<table width="778" border="1" cellspacing="0" cellpadding="1"
bordercolor="#0000FF">
<tr>
<td>50K 内容</td>
</tr>
</table>
</td>
</tr>
<tr>
<td><table width="778" border="1" cellspacing="0" cellpadding="1"
bordercolor="#0000FF">
<tr>
<td>50K 内容</td>
</tr>
</table></td>
</tr>
<tr>
<td>
<table width="778" border="1" cellspacing="0" cellpadding="1"
bordercolor="#0000FF">
<tr>
<td>50K 内容</td>
</tr>
</table>
</td>
</tr>
```

```
</table>
```

如果让表格之间保持独立，即把所有的小表格及内容从大表格中抽离出来，则页面内容就会分散在 3 个小表格里，每个表格里内容的代码量在 50KB 左右，如下面代码所示。这时，浏览器只须下载并执行 50KB 左右的代码量就可以显示相应的内容，如图 5-1 所示。

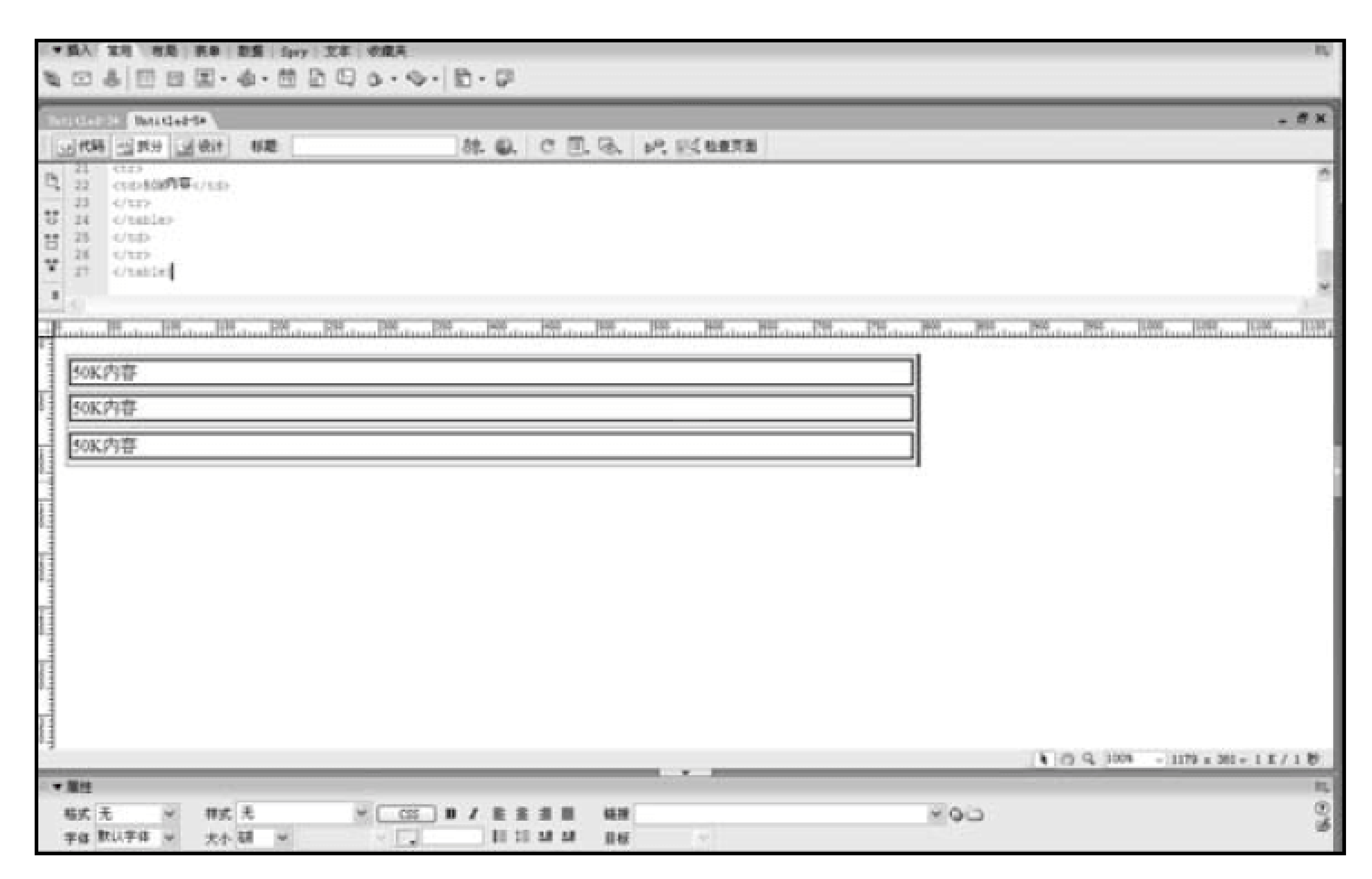


图 5-1 表格间相互嵌套的页面

```
<table width="778" border="1" cellspacing="0" cellpadding="1"
bordercolor="#0000FF">
<tr>
<td>50K 内容</td>
</tr>
</table><br>
<table width="778" border="1" cellspacing="0" cellpadding="1"
bordercolor="#0000FF">
<tr>
<td>50K 内容</td>
</tr>
</table><br>
<table width="778" border="1" cellspacing="0" cellpadding="1"
bordercolor="#0000FF">
<tr>
<td>50K 内容</td>
</tr>
</table>
```

经过精简代码，有效地删除了页面中的“冗余”代码，降低了页面体积，提高了页面的搜索引擎友好性。

尽管我们总是想方设法地降低页面的体积，但并不代表页面的体积越小就越好。因为页面体积越小意味着页面中的内容就会越少，能向用户提供的信息量也就越少。然而，页面的体积究竟多大才合理呢？这取决于网站的类型。一般来说，页面的体积控制在 100KB 左右会

更受用户及搜索引擎的欢迎（仅指 HTML 源代码的部分，不包括图片及其他多媒体元素）。

如果一个页面超过 100KB，搜索引擎就可能会拒绝收录 100KB 后面部分的内容，甚至放弃抓取该页面。

除非这个页面具有更高的权重或者相对重要的内容，如网站的首页，即使有三四百兆字节，也极有可能会被收录（如新浪网的首页）。由此得到以下结论：

- *一个页面的体积保持在 100KB 左右对于搜索引擎最为友好。*
- *搜索引擎更重视页面中前 100KB 的内容。*
- *搜索引擎更愿意收录体积在 100KB 左右的页面。换句话说，如果网站上大部分的页面保持在 100KB 左右，则被搜索引擎收录的机会就会大大增加。*

随着各方面技术的成熟，搜索引擎会慢慢接受 200KB 甚至体积更大的页面。但是，为了提高用户体验，在此建议在规划页面的时候，页面体积要尽量保持在 100KB 左右。

经过上面的操作，我们基本上清除了页面中的冗余代码。接下来，将介绍代码优化的另一个重要任务，即提高页面相关性，主要包括页面头部优化及权重标签的使用。

5.3 标头优化

HTML 代码中<head>…</head>之间的区域成为网站的标头，通常包括页面的标题、描述及关键词等。

大家都知道在网站中 title 标题关键词是最为重要的，然而在标签中 title 标签也是最重要的，所以说一个好的标题可以很简洁有效地概括出一个网站的内容。那么对于标题的布局而言，网名就一定要采用完整匹配，也可以采用分词匹配，这样的效果在标题优化中是很不错的。

1. Description

它在网站担任的职责就类似于人们在应聘工作时候的自我介绍，并且可以给用户提供一些内容上面阅读的重要标签之一。它的作用是概括站点内容让用户了解，并且用户在搜索引擎搜索某些关键词的时候可以看到你网站域名上方的一段描述，也是给搜素引擎一个参考，还可以作为引导 se 和 ue 的一个重要路标。所以充分地体现出 Description 标签的重要性，但是我们在描述的时候一定要简单明了，字数一定要在搜索结果范围之内，如图 5-2 所示。

123宠物网—中国最专业的狗主人社区及宠物狗交易市场
123宠物网是中国宠物行业最具专业的宠物狗主人社区平台，中国最具专业的宠物狗交易市场，在这里您可以交流养宠物狗狗的心得，在论坛、家园结交狗友并分享学习养宠物...
www.123pet.com.cn/ 2012-4-20 - 百度快照

图 5-2　123 宠物网描述

2. Keywords 密度

因为对于不同的搜索引擎而言 Keywords 密度也是不一样的，所以这点在优化中往往都被忽略，并且也不在作为搜索引擎给予一个好排名的重要因素，只要大家合理地掌握好关键词的密度，一般都不会因此造成不必要的祸端，所以笔者不多给予介绍。

3. <strong>、<em>和<u>标签

对于网站而言这三种标签的作用非常类似，它们都是对网页内容中的重要关键词来进行加强突出，让搜索引擎更为重视从而达到目的。在此笔者就不多予介绍，另外就是<b>、<i>等标签，很多的站长认为它们同样也可以达到突出加强关键词的效果，然而这样的观念是错误的，所以笔者还是建议大家使用<strong>、<em>和<u>这三种标签来进行加强网站关键词。

5.4 权重标签的使用

在搜索引擎算法中会默认一些标签，对关键词加权，也就是为关键词加分，对提升关键词的排名有很大的效果。在关键词优化过程中合理地使用这些权重标签，不论是对读者还是搜索引擎，都是相当有好处的。

在 SEO 中常见的权重标签<h>：这个在站长优化标题关键词的时候用得最多，也是最为常见的，效果也是不错的。大家都知道在<h>标签当中权重依次为<h1>、<h2>、<h3>……。它的好处就是便于搜索引擎快速地掌握内容的大概含义，从而凸显内容中最为重要的关键词，然后授予最高的权重。但是我们在做<h>标签的时候一定要注意切勿把<h>标签用在 LOGO 上面，还要注意在页面中出现的次数不宜过多。

其他常见的权重标签有：<strong>、<em>、<b>、<i>和<big>。

1. <strong>标签

<strong>标签是权重标签中的代表，在搜索引擎中能够得到高度重视。将<strong>应用到你想强调的文本或者关键词中，这部分内容将会得到读者和搜索引擎的重视。从 SEO 的角度来说，<strong>标签能够突出关键词，表现重要的内容。<strong>标签在浏览器中的表现为加粗。

2. <em>标签

<em>标签在浏览器中表现为斜体，表示的意思也是强调主体内容，吸引用户关注。它在权重标签中的重要度仅仅次于<strong>。当然，这个斜体一般应用于强调内容，而不是强调关键词。

3. <b>标签

<b>标签在 Inertnet 和 Netscape Communicator Web 等主流浏览器中都是支持的，它在浏

览器中表示的效果和<strong>是一样，都是加粗，但是他们表达的意义完全不同。<b>标签在本质上没有多大的逻辑意义，通常被用来强调页面的作者。而在 SEO 中，<b>标签有时候会破坏网站的权重，很多时候使用它会起到反作用，不合理利用会带来很多负面效果，所以，在关键词优化过程中不推荐使用。

4. <i>标签

<i>标签在浏览器中的显示效果为斜体，它往往被用来标注内容，突出内容要点，实际上，合理利用斜体标签可以突出内容的重点性，但是不能刻意去针对关键词，否则也会给网站的优化带来负面影响。

5. <big>标签

<big>标签是 HTML 中比较老的一个标签了，在 CSS 中基本已经被淘汰了。不过，在搜索引擎中仍然是一个权重标签，它能够引起搜索引擎的重视，突出关键词，增加关键词的权重。不过，在这里不推荐使用。

值得注意的是，在使用各种权重标签的过程中，重点要注意合理使用，不能针对 SEO 来使用某个标签，否则不仅不能带来权重，而且会破坏网页的权重，让搜索引擎反感，得不偿失。

5.5 图片优化

123 宠物网针对目前搜索引擎对图片的抓取规律及图片 SEO 对网站排名的效果，现针对图片优化做以下解释并就图片 SEO 做出具体的优化执行方案，如图 5-3 所示。

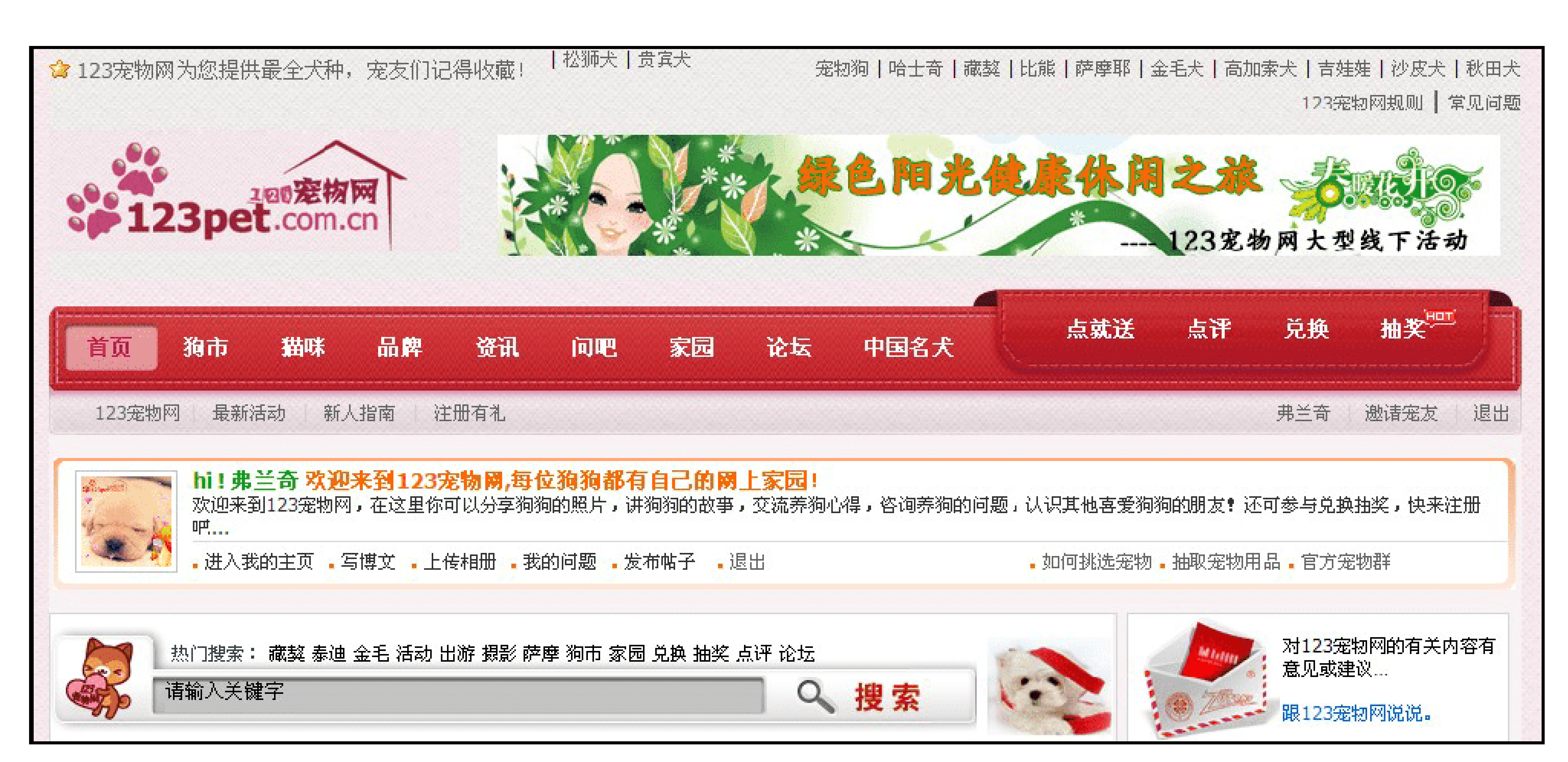


图 5-3　123 宠物网图片策略

- 百度图片搜索认为是否和关键词有关联的依据是 alt。
- 百度图片分析图片新闻和关键词的相似程度的依据是 title+alt。
- 百度图片搜索认为有较高价值的图片新闻的依据是 title+alt+keywords。

图片新闻的特点如下。

- alt=搜索关键词或基本相同（75%）。
- alt=tilte 意思是，title 里可能包含了 alt，但搜索关键词，也可能没有 alt（75%）。
- 经过分析 keywords 基本与 alt 无关（80%），很明显的可以看出 alt 在图片搜索中发挥了非常重要的作用，虽然表面上看似很容易被理解，title+alt+keywords 三者越相似，越容易被收录，排名越靠前。

因此网站图片 SEO 应该注意做到以下细节的优化处理。

1. alt 元素

这是优化网站图片最重要的一点，即 alt 部分的处理。

使用 alt 属性是为了给那些不能看到你文档中图像的访问者提供文字说明。除此之外你还可以添加一些特殊的标注，如图片说明、评论、设计群体、位置和主题。搜索引擎在抓取内容时也会看图片周围的文字来决定图片的相关性。标注旁边的文字和锚文本旁边的文字对图像搜索排名有一定的影响。

如果你对图片的标签和文字说明有兴趣的话也可以试一下。如 imgsrc="promote.yidaba.jpg/"alt="123 宠物狗市"，但是这个代码组合还不算完善，一个标准的图片嵌入代码应当包含哪些元素呢？从 Pearsonified 那里我们得知，它应当包含 5 个关键性元素 src、width、height、alt 和 title。其中 alt 元素最为关键。

如图片，长宽为 485×309。那么它规范的 HTML 代码就应当为：imgsrc="123pet.com/"width="485"height="309"alt="图片 seo"title="seo 图片优化策略及执行方案" />。

2. 图片文件名

图片的命名，取一个和文档相匹配的图像名称很重要。图像名称将会出现在图表图像的搜索结果中，当搜索用户在浏览他们喜欢的图表图片中，还可以和用户形成沟通。不要以为图像编辑器会对你的默认图片名称进行优化，因为默认的名字对搜索引擎没有任何的沟通作用，因此应该事先设定一套命名结构体系。

建议，如果你有一张 123 宠物网狗市的图片，那么就应当使用像 123gous.jpg 这样具有描述性质的文件名，而不是使用 123.jpg 这类只是代号的名字，当然也可以使用文字作为图片的文件名。

3. 图片标题

这点和网页优化中的 title 标题一样，是张什么内容的图片，就给这张图片取个相符的标题。如果你用 Google 搜索图片会发现每张图片下都会有一段描述性文字，其中关键字用粗体

显示。这段描述行文字一般都是位于原文章图片周围。

建议，在图片周围添加包含关键字的描述信息，将可以优化网站图片在搜索引擎中的排名，如图 5-4 所示。



图 5-4　123 宠物网狗市频道

4. 图片文字介绍

排在前面的图片，基本都是关于这张图片的一个专栏介绍，以告诉搜索引擎这个图片的重要性。

5. 外部链接

网站优化需要外部链接，同样，图片的 SEO 优化也是需要外部链接，链接始终是网络最根本的基础，最优秀的图片才会在网络上蔓延，搜索引擎也就是分析这张图片的外部链接，从而来判断这张图片是不是受欢迎。

使用超链接打开图片，那么超链接的名字应当包含图片的关键字。

如你要链接一张“123 宠物网”的图片，就不宜使用“点击获得完整尺寸”等链接文字，而应当试着使用诸如“123 宠物网”之类的命名形式。

6. 图片的存储

在网站图片使用过程中，应谨记使用自己的而不是其他网站的图片。因为自己最了解网站的速度及相关路径和文件设置，如果是引用外部链接打开图片的话，很可能因为外部网站的服务器问题而导致图片无法打开。

7. 图片的尺寸

如你链接的是一张网站优化工具的图片截图，那么图片的大小就应当符合规范，这样才不会因为设置不规范使图片变形而导致无法查看。

8. 检测搜索索引

为了检测我们对网站图片的优化方法是否合适，可以对 Google 检索的网站图片进行检测。

9. 两个关键词的组合搜索的结果

图片中是以宠物品种的名字命名的，如金毛犬。alt 时有的人喜欢直接写金毛犬图片，这个在只有一个图片的情况下是可以的。如果有两张图片，或者更多的时候呢，第一张 alt 写“金毛犬的图片”有人喜欢这样搜索，反正搜索金毛犬图片的时候，这个 alt 是包含了的。第二张图片，alt 写“金毛犬的照片”，这就是说在图片的 alt 中写上和这张图片有利于搜索到的关键词。

也就是我们所说的关键词匹配，这里讲的是宠物种类，而其他种类也可以利用这个方法，加上图片的相关词（只要具有关联的就可以加），当然也可以无中生有。

5.6 沙盒效应

很多刚刚上线的新站在 Google 很难有好的排名，或者根本没有排名，这种现象我们称之为“沙盒效应”。为什么产生这种现象，以下是笔者的个人看法：以前大家发现堆砌关键字排名很好，所以一窝蜂地堆砌关键字，Google 就调整了算法，所以堆砌关键字就不管用了；后来大家又发现外链接很好用，就给新站加许多外链让排名上升。但 Google 算法中有很重要一部分是基于 PR 的，不会过多地对外链降权。这样，Google 就弄出了这个沙盒，把它们看起来可疑的网站放到沙盒里面考察，一段时间内加任何外链是不起作用的，这样就大大降低了大家加外链的兴趣，这也是 Google 对付 SEO 的一个方法。

沙盒英文单词叫 Sandbox，是 Google 反击垃圾网站的重要措施，出现在 2004 年 3 月，也就是举世瞩目的 Google 弗罗里达风暴和奥斯丁风暴之后的半年。有了沙盒，Google 仍然像过去一样迅速收录网站，从互联网上从新的网站里努力获取新鲜的信息，但是不再像以前那样信任这些新网站。所有新收录的网站都要经过“适用期”或者“观察期”。经过时间的考验，如果这些网站能顺利通过 Google 多次的反“泛滥”过滤器（spam filter）的考验，最终这些网站就被放回搜索的“大海”之中，参与正常的排名竞赛之中。这个过程少则两、三个月，长则要一年。新网站可以采取积极的措施，争取早日走出沙盒。

1. 如何摆脱沙盒效应

在最重要的关键字搜索排名前列的数日风光之后，许多新网站就会被雪藏于 Google 搜索结果底部，好像根本不存在一样。即使该网站可能具有很高的 Google 网页级别（PR 值），拥有很多强有力并且主题相关的引入链接和丰富的内容，它还是会遭遇到令人沮丧的“沙盒效应（Sandbox）”的影响。

你需要考虑怎样才能从沙盒效应中摆脱出来。当抑制因素排除时，你在沙盒惩罚期间所做的工作可以使你的网站重新出现在搜索结果中。

沙盒效应作为 Google 使用的搜索排名阻尼过滤器（search ranking damping filter），是用

于网站首次发布后的前两个月至4个月内获得“新站奖励”（fresh site bonus）的网站。这个给予新网站的奖励是指短时间内在搜索结果中获得非常高的排名，因为Google偏爱新内容。在沙盒期间，新站应该不断完善各方面SEO，沙盒效应过后，一般会有不错的排名。

由于原来的新内容变得稍微陈旧，沙盒过滤器开始起作用，这就是沙盒效应。一个网站陷入沙盒的平均时间长度大约为90天，但是很少有网站能在沙盒中持续待4个月。

大多数网站按照关键字的类型分享阻尼下降效应（damping down effect）。沙盒过滤器适用于所有网站，而不管它们是否有很多引入链接，即使相关性很强的链接也没什么区别。沙盒效应必然是Google的算法之一，因为它很普遍。

2. 什么情况会进入沙盒

当网站的主要关键词竞争性比较强时，更可能出现沙盒效应；对于竞争比较小的企业站，个人小站，或者公益性网站等，出现这种现象的机率很小。当你的小站链接不是自然增长，而是突然增加了大量的外部链接，也可能会被Google收进沙盒观察，因为有些黑帽SEO通常靠购买大量链接制造垃圾网站。

3. 沙盒效应的意图

很多观察者都认为，沙盒过滤器的目的是劝阻不择手段的Web站点管理员不要使用违反Google站点管理员指南的做法。Google试图瓦解一些手法，如使用垃圾站点建立初期的流量，购买过期的域名以获得其之前存在的Google网页等级作为跳跃的起点。

除非建立更长久的链接，否则短期的链接租用和放置对一个新网站来说并没有什么好处。也存在这种可能性，即Google在索引一个网站后的最初几个月中，并不会给予其全部的网页等级（PageRank）。阻尼效应（dampening effect）造成的缺乏信用的网页等级（PageRank）会降低过期域名引入链接的价值。

从这种情况可以推断出Google是否试图在阻止普遍的购买过期域名的行为。如Google工具栏可能显示为PR7，但是算法可能赋予其网站搜索排名的PR值为零。当然，这种情况下也有可能导致相当高的Google排名。

Google可能打击的另一个方面是垃圾站点。如果一个纯粹的垃圾站点制造者（spammer）在网站运行后的前几个月中不能获得好的排名，则可能会关闭他们的垃圾站点。打击spammer是Google的长期目标。

然而，Google方面声称的任何良好的意图都间接打击了那些没有违反任何规则的网站。许多Web站点管理员，尤其是那些完全不知道SEO整体情况的站长，通常对于所发生的事情非常困惑。如他们在Yahoo！和MSN的搜索结果中排名很好，但却不知道为什么不在Google的索引目录中。还有很多人错误地认为他们无意中触发了Google的惩罚。

Google打击的另一个行为是购买链接获得初始的网页等级（PageRank）。Google公司的人可能认为一个网站应该随着时间的过去而自然获得链接。他们认为购买的链接不是网站自然发展的表现。但是，只有少数几个自然链接的网站，也遭遇了同样的沙盒阻尼效应（dampening effect）。

这种情况也完全有可能发生，即一个网站没有被放在沙盒中，但是它的链接正在被监测。算法可能会考虑其链接的时间，它们的链接站点，Internet 服务商拥有的链接范围，以及通常情况下整个链接的多样性。

4. 沙盒效应不一定是坏事

如果你的网站正好进入沙盒，最好不要惊慌。Google 并没有惩罚你的网站。你也知道你的网站已经被 Google 索引。相反，这只是一个正常的现象，是 Google 算法有些令人困惑的和苦恼的地方。如果网站由于“新站奖励效果”而在搜索结果中获得很高的排名，则很有可能会尝到苦果。

5.7 小结

SEO 其实就是如何把网站做得更好，把不好的东西去除或者改善。网站页面的优化，也是一项相当重要的基础优化工作。在网站的运营当中，应该随时注意每一个细节。同时，大家在对网站进行更新时应该注意所添加的页面内容的篇幅，正文是否能够足以阐明页面主题。

第 6 章　网页组成结构

要做好网站的内部优化就需要了解网站的组成结构，需要对网站整体构架有个全面的认识。本章主要介绍网页的组成结构及对优化的影响，包括网页结构简介的基础知识，网页组成的基本概念。

6.1　网页结构简介

网页的构架就是网页的内容格局，是创建制作页面的重点之一。合理的网站构架有助于提高搜索引擎收录的数量以及用户的体验度。

6.2　网页组成

设计网页时要组织好页面的基本元素，同时再配合特效，即可构成一个绚丽多彩的网页。网页的组成对象包括文本、图像和超级链接。网页主要是由导航栏、栏目及正文，这三大元素组成。网页结构的创建、网页内容的布局与规划其实都是围绕这几点而展开的。

1. 导航栏

导航栏是指位于页眉区域的，在页眉横幅图片上边或下边的一排水平导航按钮，它起着链接博客的各个页面的作用，如图 6-1 所示。

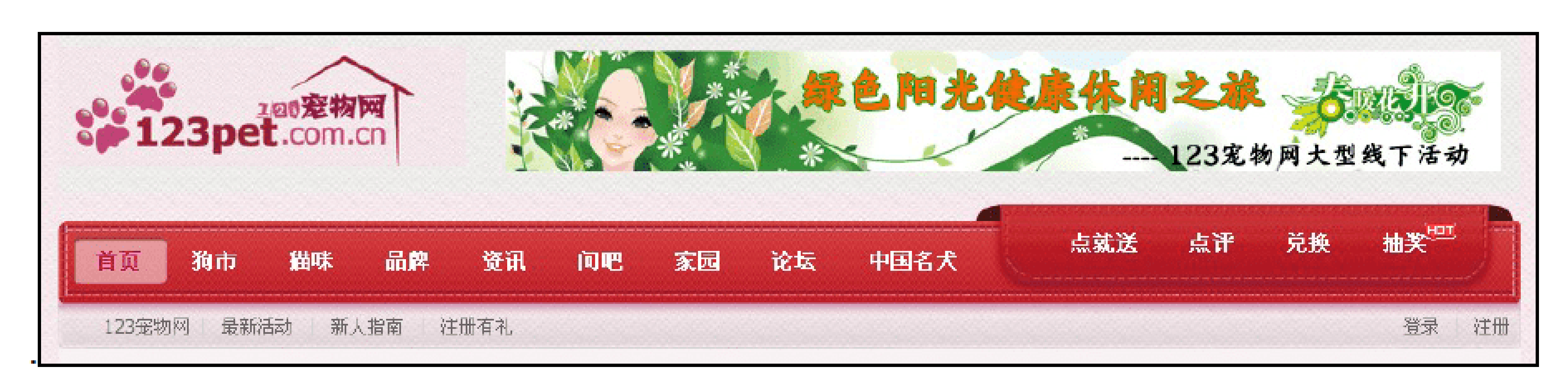


图 6-1　123 宠物网导航栏

通常情况下，WordPress 平台主题里的导航栏是不允许将其链接任意定制成其他链接

的，但是每当你撰写了一个新页面，导航栏就会自动多出一个指向该页面的按钮链接。网站使用导航栏是为了让访问者更加清晰明了地找到所需要的资源区域，以便寻找资源。

2. 栏目

栏目是指页面中存放相同性质内容的区域，在对页面内容进行布局时，把性质相同的内容安排在页面上的同一区域，可以帮助用户快速获取所需信息，对网站内容起到非常好的导航作用。

做好网站栏目的构架能够很直观的看到整个网站的结构、涵盖的内容，也是对整个网站宏观的一个定位。像“123 宠物网”这样的大型网站栏目都非常全面，因为它们的用户非常多。而中小型网站则偏重于细分化，专论某一行业或某一细分领域，但是网站栏目的策划构建同样非常重要，如图 6-2、图 6-3 所示。



图 6-2 新闻栏目



图 6-3 狗市栏目

3. 正文

正文是指页面中的主体内容。例如一个文章类页面，正文内容就是文章本身。对于展示产品的网站，正文内容就是产品信息，其中包含文本和图像。

（1）文本

文字是网页的主体，虽然利用 Flash、图形文字也可以达到同样的效果，甚至超出纯文本效果，但是网页文字的优势还是无法被取代。因为纯文本所占用的存储空间非常小。然而在页面上用同样的字体显示，会使页面过于呆板。在页面中适当调整文字的大小、颜色等，也可以改善页面的效果。

①标题

在一个网站的网页或一篇独立的文章中，通常都会有一个醒目的标题。告诉访问者这个网站的名字或该文章的主题。而现在很多网页设计者用图像或者 Flash 动画来代替文字标题。

②文字大小

适当地调整文字大小，可以使文字编排得更加生动活泼，达到更好的效果。

③段落

文章段落分明，有层次感，才能让访问者更好的阅读，也使得页面看起来整洁、美观、大方。

④样式

粗体、斜体和下划线是最基本的文字样式，然而在网页中不可过多地应用下划线，那样会使访问者误以为是超级链接。在页面的适当位置添加一些字体样式，会使页面更具有观赏性。

⑤字体颜色

也可以为页面上的字体添加颜色，以强调页面中的重点，使页面也变得绚丽起来。但是一定要注意颜色的搭配。在页面上也不能过多地应用颜色，太过华丽反而会引起访问者的反感。

（2）图像

图像给人的视觉效果要比文字强烈得多。灵活应用图像，在网页中可以起到点缀的效果。但是运用不当，会使网页变得凌乱不堪。Web 页上的图像文字大部分都是使用 JPEG 和 GIF 格式。因为它们除了具有压缩比例外，还具有跨平台特性。无论访问者使用什么样的操作系统，都能够显示这两种图片。图像在网页中的应用主要有以下几种形式。

①图像标题

一般在网页中都要有标题，用以提示访问者这个网站是做什么的，从而起到导航的作用，应用图像标题可以使网页更加美观。

②背景

图像的另一个重要应用是作为网页的背景。网页背景千变万化，特别是一些个人网站，应用图片背景比较多。

③网页主图

网页上除了用小图像美化网页外，通常还会在网页上用一些较大的图片来突出网页主题，占满网页的整个空间。特别是主页中用主图的比较多，或者直接用图片取代文字作为超级链接按钮，可以使网页更加美观。

6.3　重要区域分布规律

搜索引擎对页面中每个区域的重视程度是不一样的，同样的内容出现在页面中不同的区域，所起的作用也会不同。

如“头条新闻”出现在页面的顶板或者底部，所起的作用就是完全不一样的。从用户的立场出发，普通用户在查看页面时是自上而下、自左向右地进行的。因此，页面中各区域的重要性关系是：“左上→右上→左→右→左下→右下”，所以应将所要重点展示的内容尽量放置在左上的位置，如图 6-4 所示。



图 6-4　“123 宠物网”论坛页面分布方式

利用搜索引擎分析主页时，在 HTML 源代码中是自上而下地进行的。因此，搜索引擎更加重视接近顶板的代码。正常的情况下，HTML 源代码自上而下对应页面中“左上→右上→左→右→左下→右下”的内容。

6.4 网页结构类型

网页结构实际上就是对导航栏、栏目及其正文内容这三大页面的基本组成元素进行辅助性的布局。页面结构布局好之后将有助于访问者更加清晰地分析，并在网站中找到的所需资料，不仅提高了搜索引擎的友好度，同时也加深了访问者对网站的记忆，让他有再次回来观看和深度观看的想法。

1. 导航型

导航型，顾名思义就是以网页导航为主的页面，根据网页的重点与次重点可分为频道导航型和内容导航型。

（1）频道导航型

频道导航型页面是以网站频道导航为主的页面，常用于频道比较多的网站首页，如导航性质的网站和目录网站。如果公司有很多种产品的类型，那么便可以设置一个频道导航，以便更好地展示公司的产品。在制作频道导航页面的时候只需要列出个频道的名称并且给其加上相应的链接即可，也可称其为锚文本，在锚文本的排列上根据大多访问者的浏览习惯应该把比较重要的放在最上面。

（2）内容导航型

内容导航型页面是以网站中相对重要的内容导航为主，内容导航型页面主要由栏目组成，常见于网站或者是频道的首页。在设置内容导航页时需要把性质或者特征相同的内容分成同一个栏目从而展示给用户，这样有助于用户轻而易举的找出所需要的内容，同样对搜索引擎有很大的帮助，可以使搜索引擎更快速的对页面进行抓取。

2. 内容型

所谓的内容型就是完全由内容组成的页面，没有其他任何信息。只是单纯的对内容进行一个完整地介绍与说明的网页。在给网站进行优化的时候内容型的页面是不能选择的，因为不管是访问者还是搜索引擎，当他们浏览到这个页面的时候就已经不知道再去哪里浏览了。访问者遇到这种情况便会离开网页重新搜索想找的信息。

3. 内容导航结合型

内容导航结合型主要的目的是展示内容所用，但是在展示内容的同时还增加了导航辅助，这样当一个访问者进入这个内容页面的时候，他不仅看到了所需要的内容，而且在他浏览完的时候，如果觉得这篇文章有用，那么就会通过内容页所设置的导航继续阅读其他的文章。对于搜索引擎来讲，在对内容抓取结束以后导航也会对其产生指引的效果，让其依然停留在本网站，让本网站被抓取的可能性大大提高，但是美中不足的就是内容导航结合型给搜

索引擎做的指引量远远不足于导航型对搜索引擎的指引量。所以哪个类型都有自己的好处与弊端。这就需要站长根据自己网站的实际情况而定了。

6.5 页面常用技术

为了让页面更美观，布局更合理，功能更强大，一个页面中通常会有多种内容。要在有限的页面空间里展示这些内容，就需要使用各种技术，如在对页面进行布局时，我们常用“隐藏层”及“框架”技术；而为了加强页面的视觉效果，有可能会使用 Flash 元素等。

6.5.1 隐藏层

使用隐藏层可以在有限的页面空间里展示更多的相对重要的内容。但是，并不是所有的搜索引擎都解析隐藏层的内容，如百度就会忽略隐藏层的内容，而 Google 则比较重视。

对于解析隐藏层的搜索引擎来说，会更加重视默认显示层的内容。为什么呢？从网站管理员的角度来说，应该会把相对重要的内容通过默认显示层优先向用户展示，而对于普通访问者来说，也希望能在第一时间看到最重要的内容。

隐藏层内容调用的方式有多种，常见的有“内部调用”和“外部调用”两种，接下来将分别介绍这两种调用方式的利弊。

1. 内部调用

内部调用是指隐藏层内容放在当前页面中，当触发相应时间时，直接从当前页面进行调用显示，这样会大大缩短隐藏层内容的相应时间，但同时也会带来如下一些负面影响。

- 隐藏层内容放在当前页面中会大大增加页面的体积，加长页面的下载时间，不管对搜索引擎还是普通用户都是极其不友好的。
- 在 HTML 源代码中，隐藏层内容占用一定的空间，从而会占据页面中相对重要的区域，如果隐藏层内容过多就会大大增加页面的体积，从而导致搜索引擎拒绝收录该页面或者忽略超过限制区域的部分内容。
- 对于个别并不关心隐藏层的搜索引擎来说，有百害无一利的效果。

在此，并不建议把隐藏层内容放到当前页面进行调用。如果非得这样做，则可以把隐藏层内的内容放在页面底部。

2. 外部调用

外部调用是指把隐藏层内容放在外部文件中，只有当时间触发以后才调用并显示。使用外部调用的方式处理，除了能在页面有限空间里展示更多相对重要的内容外，还不会改变当

前页面的体积。但是，搜索引擎对于外部调用的内容会进行单独的处理。隐藏于页面主体相关的重要内容不能放在隐藏层里。

总之，在使用隐藏层为页面进行布局的时候，要把相对重要的页面放在默认显示层里，而次要的内容放在隐藏层里，同时隐藏层的内容最好通过外部调用的方式调用。这样既可以克服隐藏层带来的负面影响，又能改变用户体验。

6.5.2　框架

框架是一个非常重要的页面布局工具。它可以把浏览器的窗口分割成若干相互独立却又有联系的区域，在每个区域中显示一个独立的网页文件，从而实现一个网页内显示多个HTML 文件的目的。使用框架的优点在于可以把不同层次的网页中的不变的内容放在一个框架中，把不断改变的文档内容放在另一个框架中。框架技术可以被广泛的应用到网站导航和文档浏览中，方便访问者对网页的浏览，减少访问者下载页面所需要的时间。

框架由框架集和框架组成。框架集就是框架的集合，它是在一个文档内一组框架结构的HTML 网页，定义了网页显示的框架数、框架的大小、载入框架的网页源和其他可定义的属性等。框架是框架集中显示的文档，每个框架实质上都是一个独立存在的 HTML 文档。

1. 创建框架集

选择预定义的框架集将自动设置创建布局所需要的所有框架集和框架，它是迅速创建基于框架布局的最简单方法。创建框架集的具体操作步骤如下：

步骤 01　选择菜单中的【文件】|【新建】命令，打开【新建文档】对话框，在对话框中选择【常规】选项卡中的【类别】|【框架集】|【上方固定，左侧嵌套】选项，如图 6-5 所示。

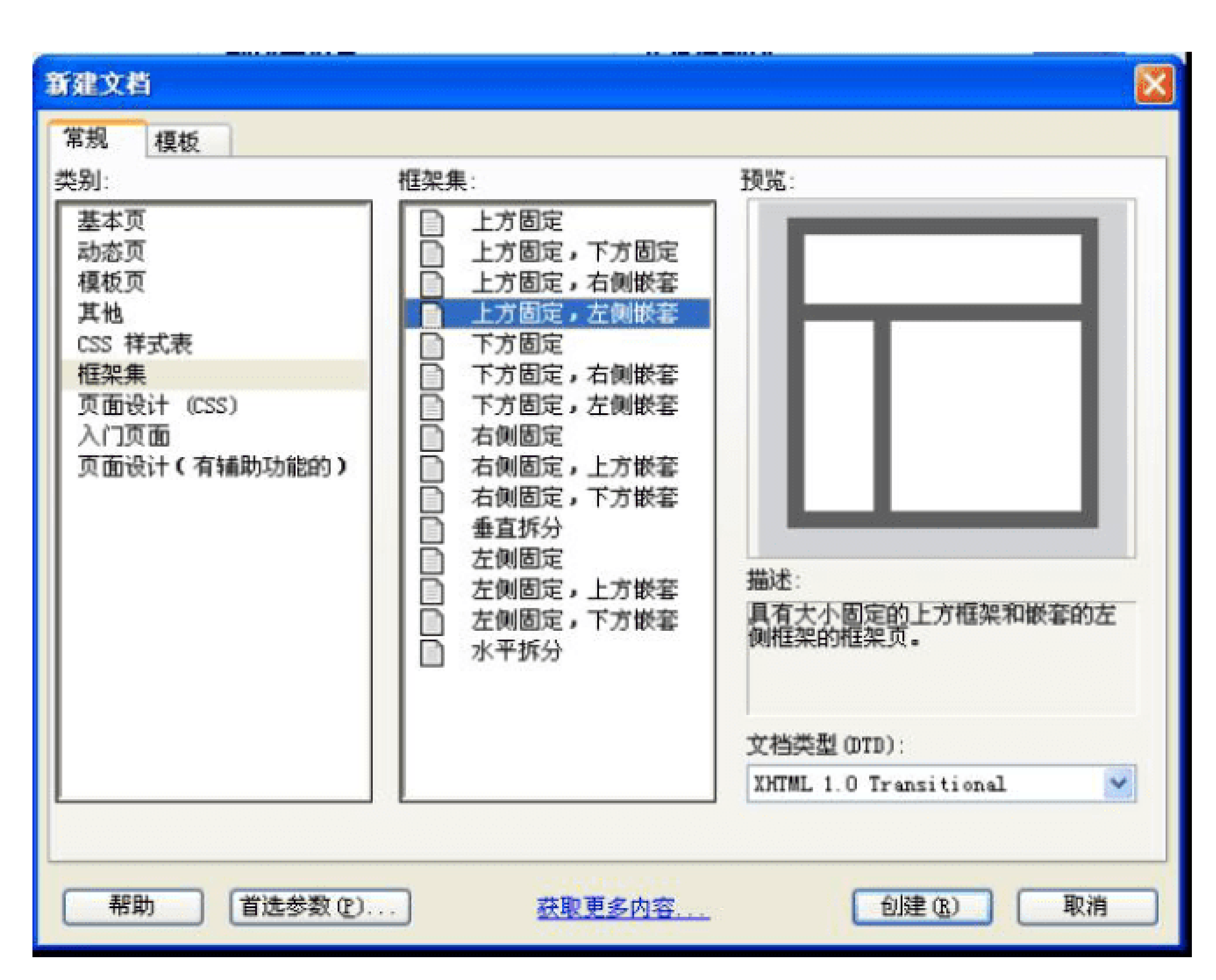
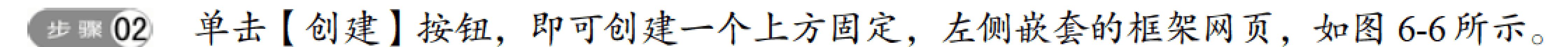


图 6-5　选择框架集



步骤 02　单击【创建】按钮，即可创建一个上方固定，左侧嵌套的框架网页，如图 6-6 所示。

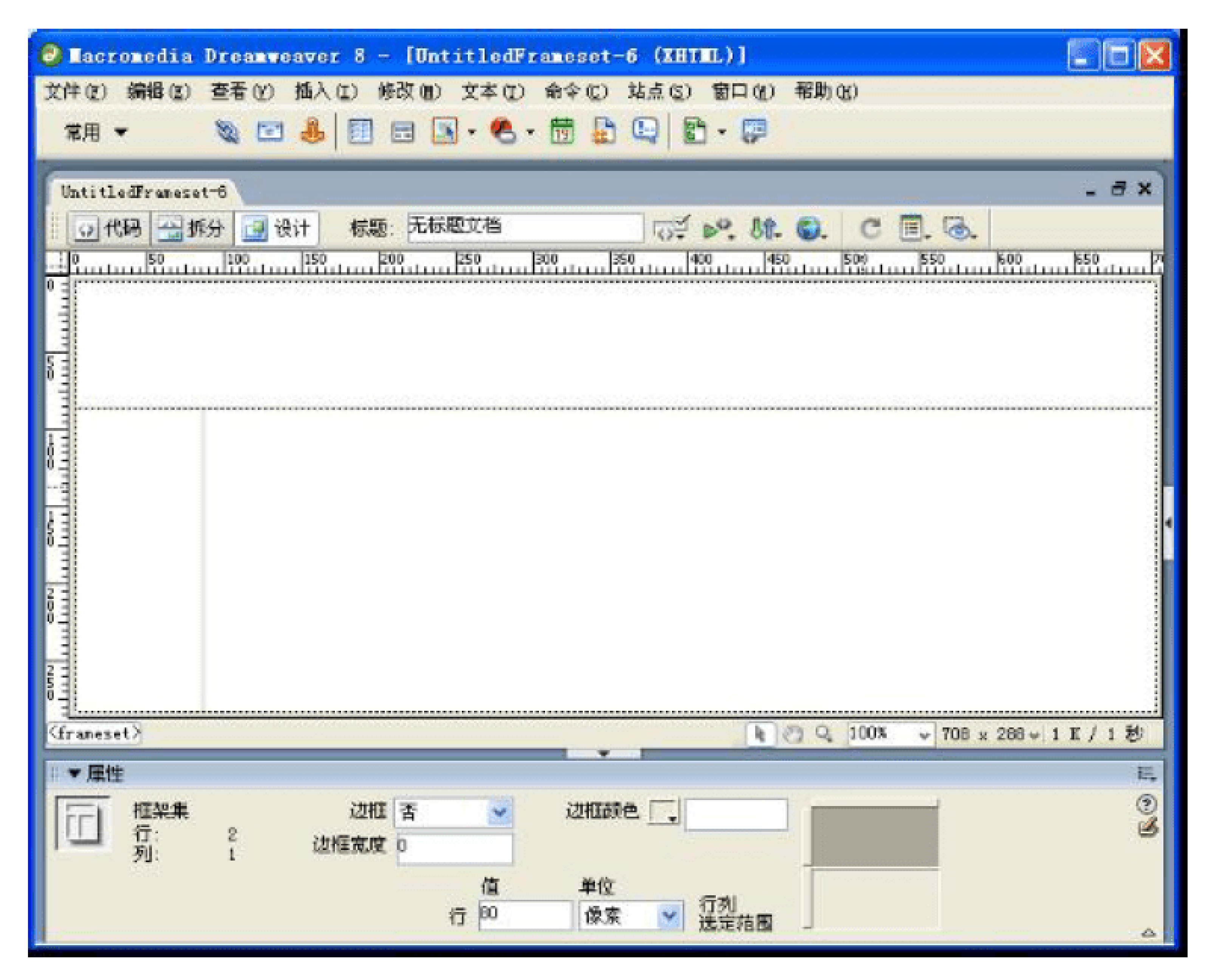


图 6-6　创建的框架

步骤 03　在【布局】插入栏中单击按钮，在弹出的菜单中选择【左侧和嵌套的下方框架】选项，即可创建框架网页，如图 6-7 所示。

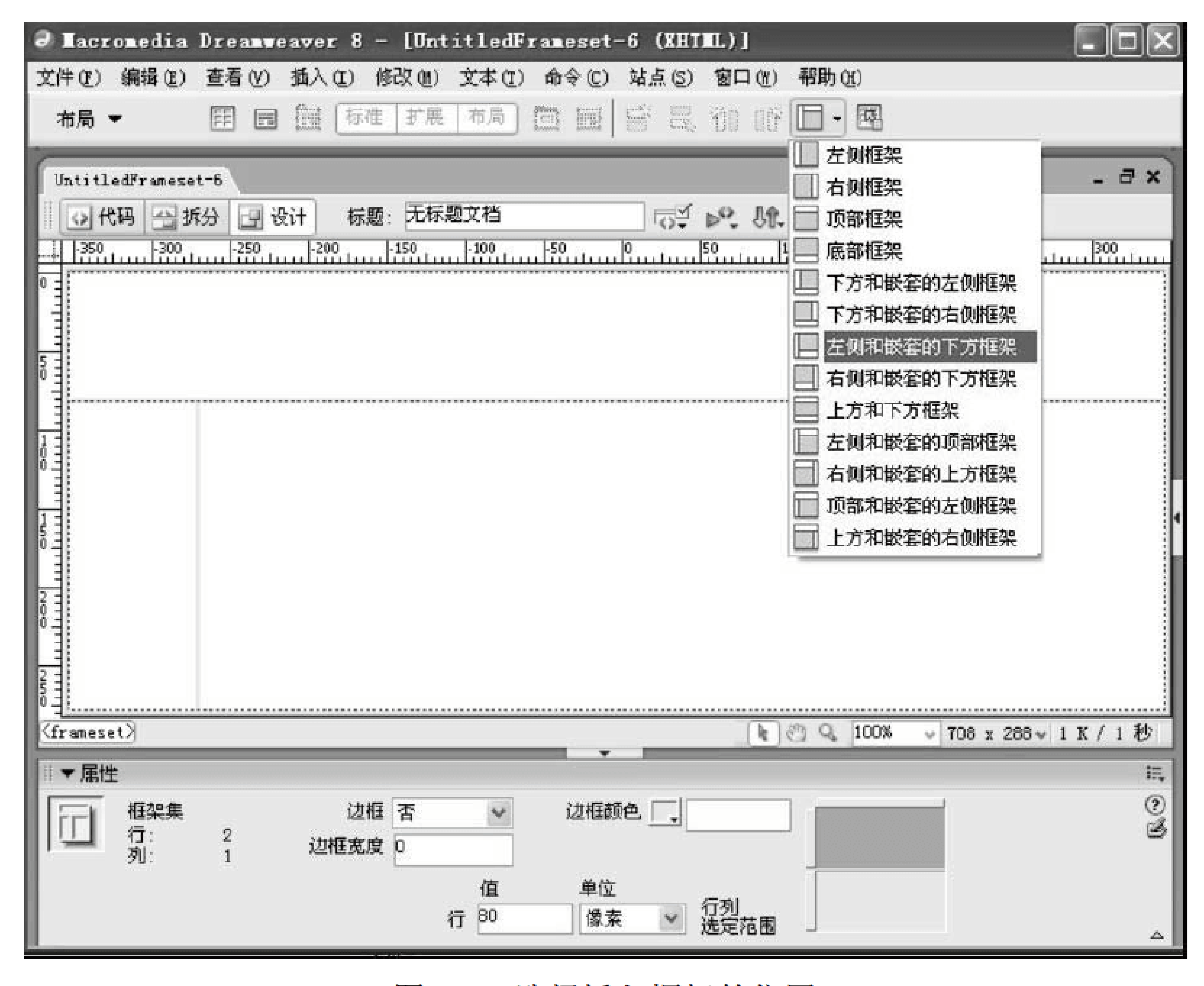


图 6-7　选择插入框架的位置

步骤04 选择菜单中的【插入】|【HTML】|【框架】|【左侧及下方嵌套】命令，即可创建框架网页。

2. 预定义框架集

插入预定义框架集的具体操作步骤如下：

步骤01 将光标放置在头部框架中，选择菜单中的【修改】|【页面属性】命令，打开【页面属性】对话框，将【左边距】设置为 0 像素，【右边距】设置为 0 像素，如图 6-8 所示。

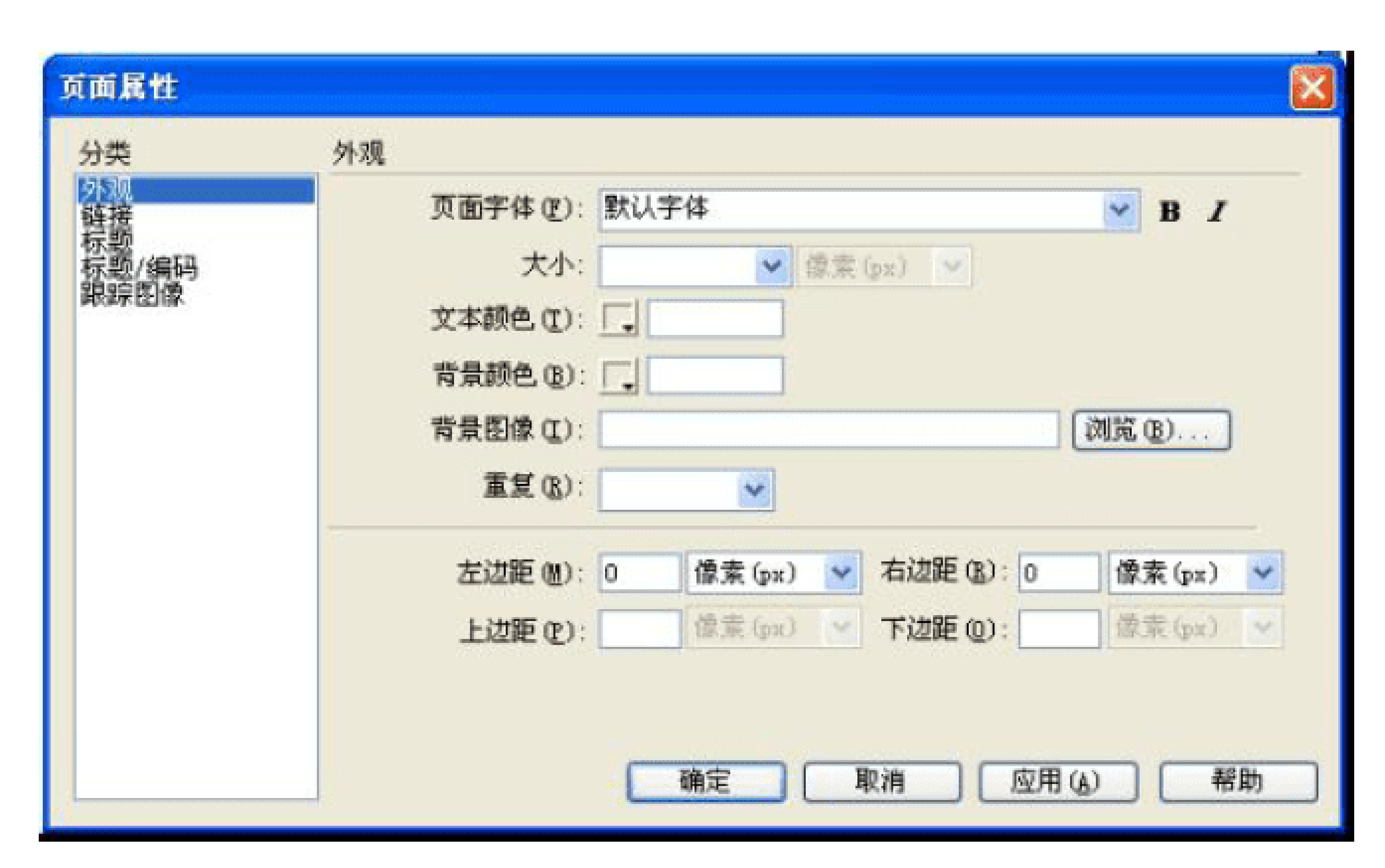


图 6-8 【页面属性】对话框

步骤02 单击【确定】按钮。选中头部框架，在【属性】面板中单击【源文件】文本框右侧的【浏览文件】按钮，打开【选择 HTML 文件】对话框，在对话框中选择 top1.html。

步骤03 单击【确定】按钮，插入文件。

步骤04 将光标放置在左侧框架中，修改页面属性，在【属性】面板中单击【源文件】文本框右侧的【浏览文件】按钮，在打开的【选择 HTML 文件】对话框中选择 leftl.html，单击【确定】按钮，插入文件。

步骤05 将光标放置在右侧框架中，修改页面属性，在【属性】面板中单击【源文件】文本框右侧的【浏览文件】按钮，在打开的【选择 HTML 文件】对话框中选择 right.html，单击【确定】按钮，插入文件。

3. 保存框架和框架集文件

在浏览器中预览框架集之前，必须保存框架集文件以及要在框架中显示的所有文档。可以单独保存每个框架集文件和带框架的文档，也可以同时保存框架集文件和框架中出现的所有文档。保存框架和框架集文件的具体操作步骤如下：

步骤01 新建一个上方固定、左侧嵌套的框架网页。选择菜单中的【文件】|【保存全部】

命令，打开【另存为】对话框，此时整个框架集内侧出现阴影，所以询问的是框架集的名称，在【文件名】文本框中输入index.html，如图6-9所示。

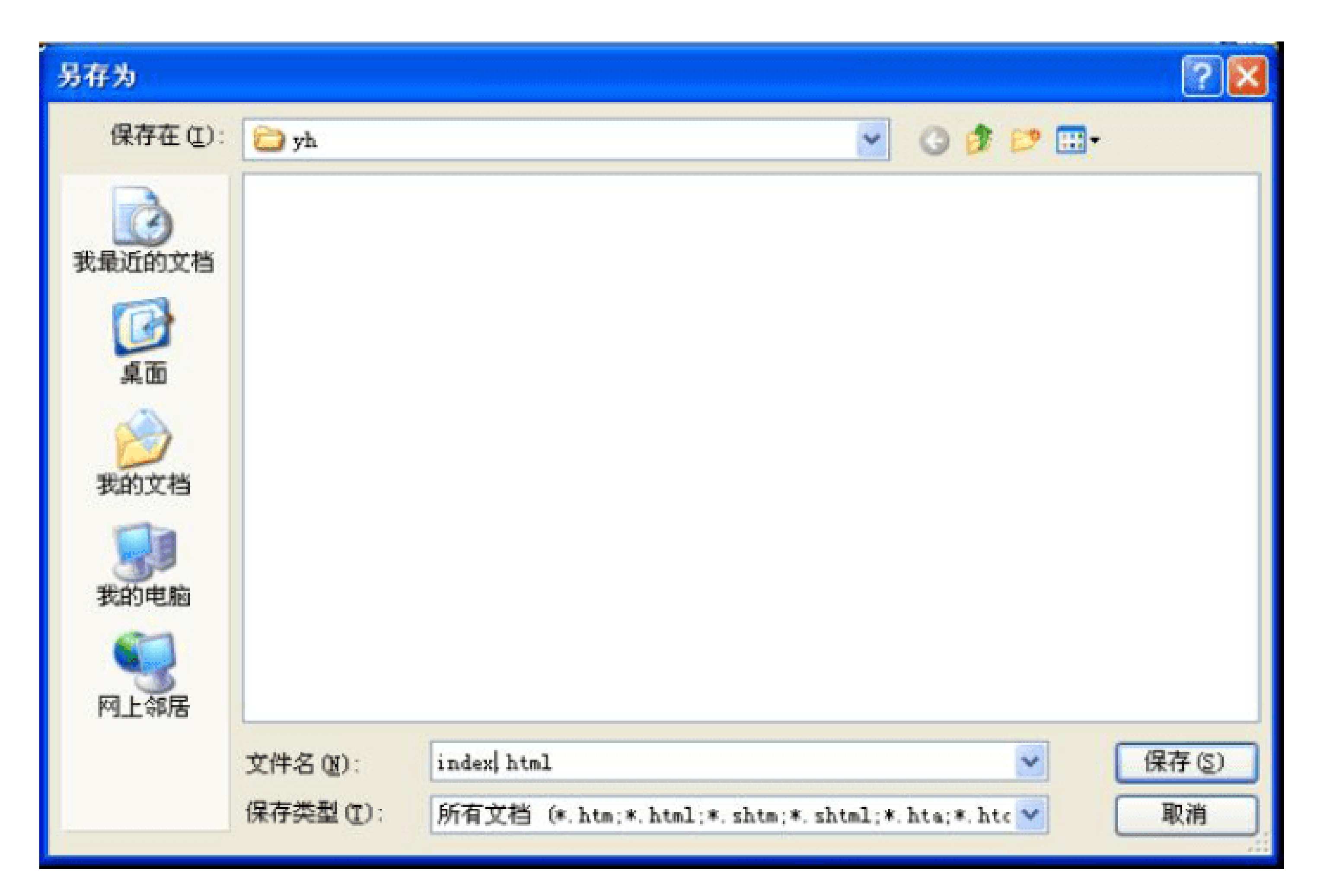


图6-9 保存框架集

步骤02 单击【保存】按钮，打开第2个【另存为】对话框，因为右侧框架内侧出现阴影，所以询问的是右侧框架的名称，在【文件名】文本框中输入right.html，如图6-10所示。

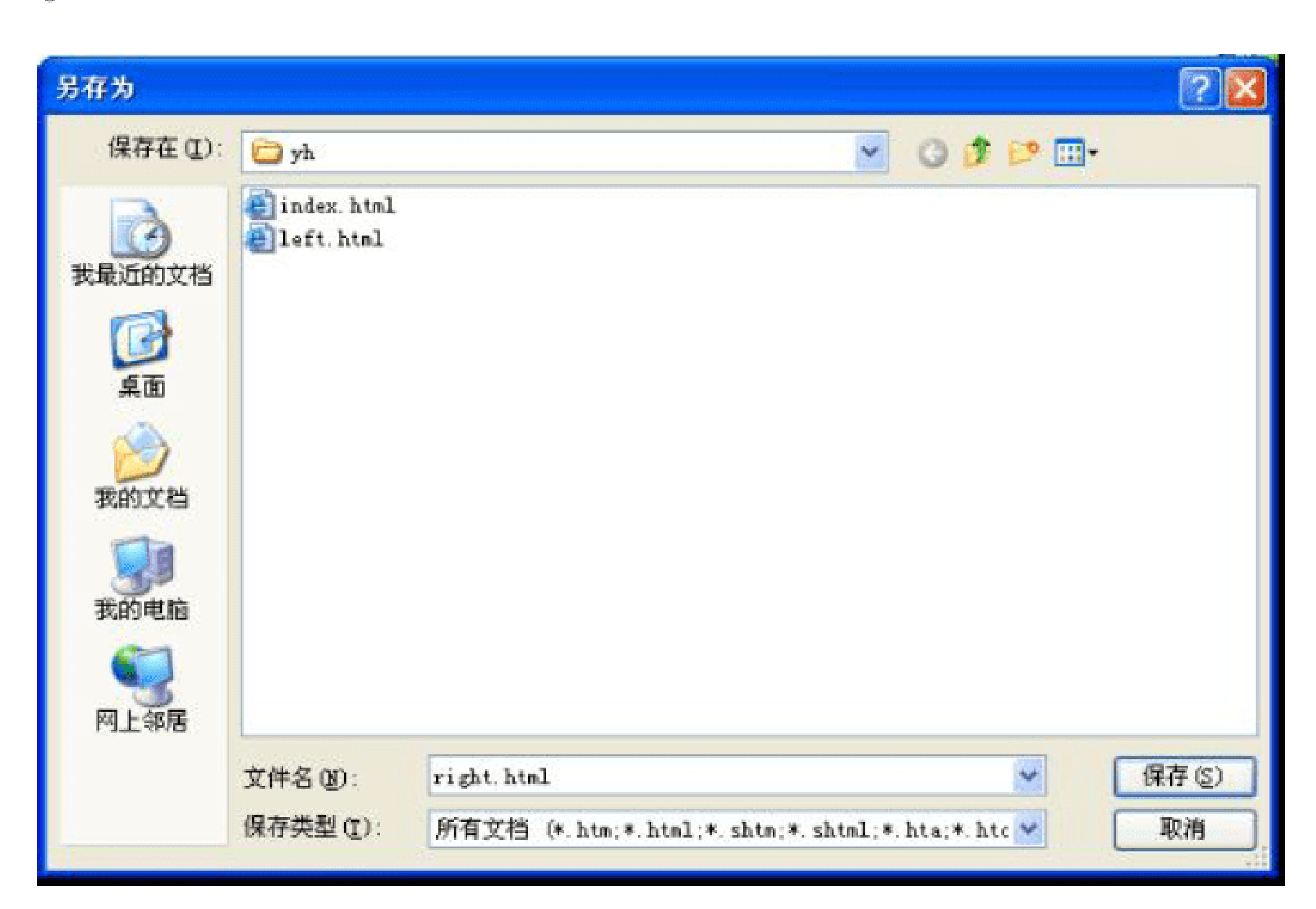


图6-10 保存右侧框架

步骤03 单击【保存】按钮，打开第3个【另存为】对话框，因为左侧框架内侧出现阴影，所以询问的是左侧框架的名称，在【文件名】文本框中输入left.html，如图6-11所示。

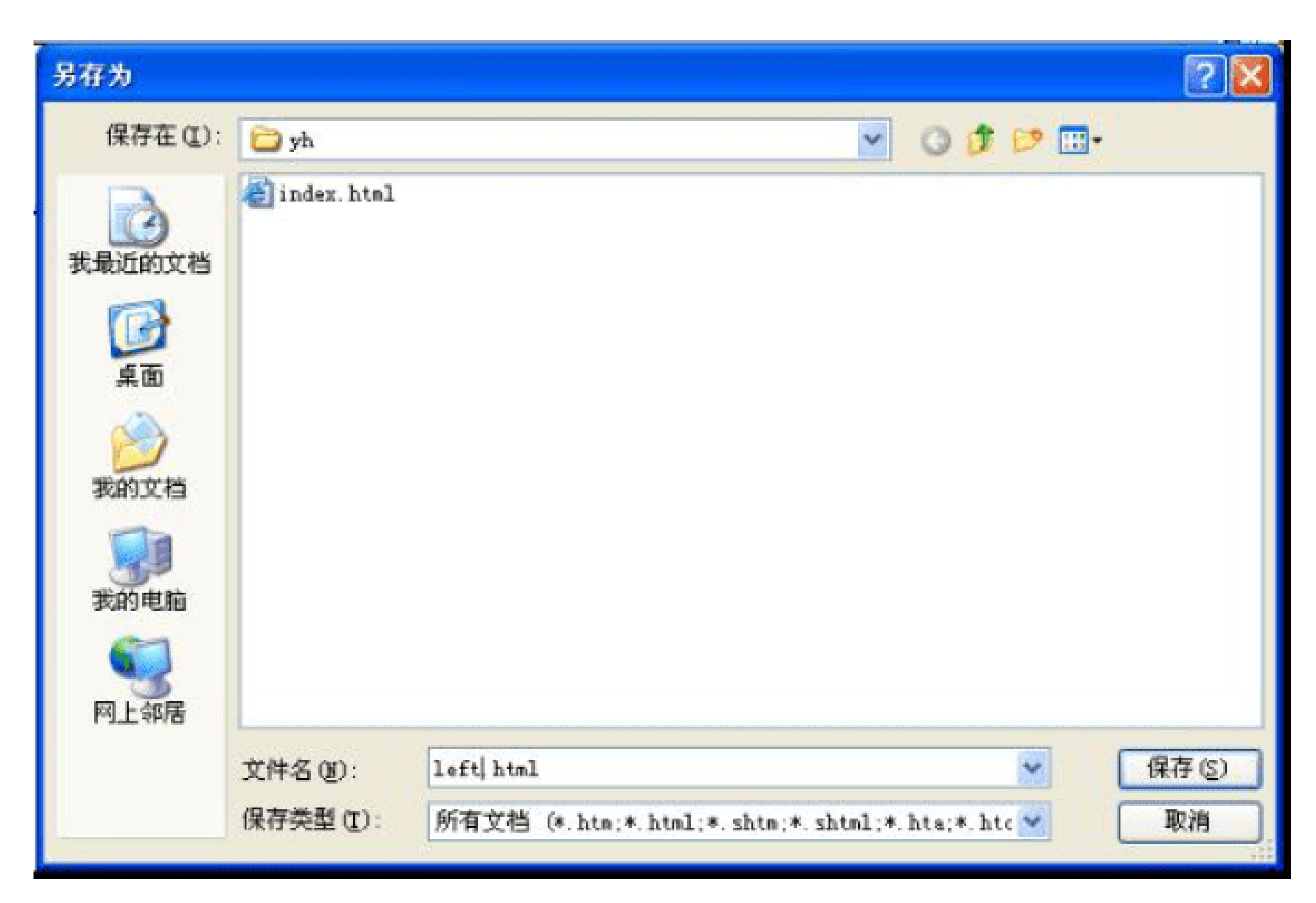


图 6-11　保存左侧框架

步骤 04　单击【保存】按钮，打开第 4 个【另存为】对话框，因为头部框架内侧出现阴影，所以询问的是头部框架的名称，在【文件名】文本框中输入 top.html，如图 6-12 所示。

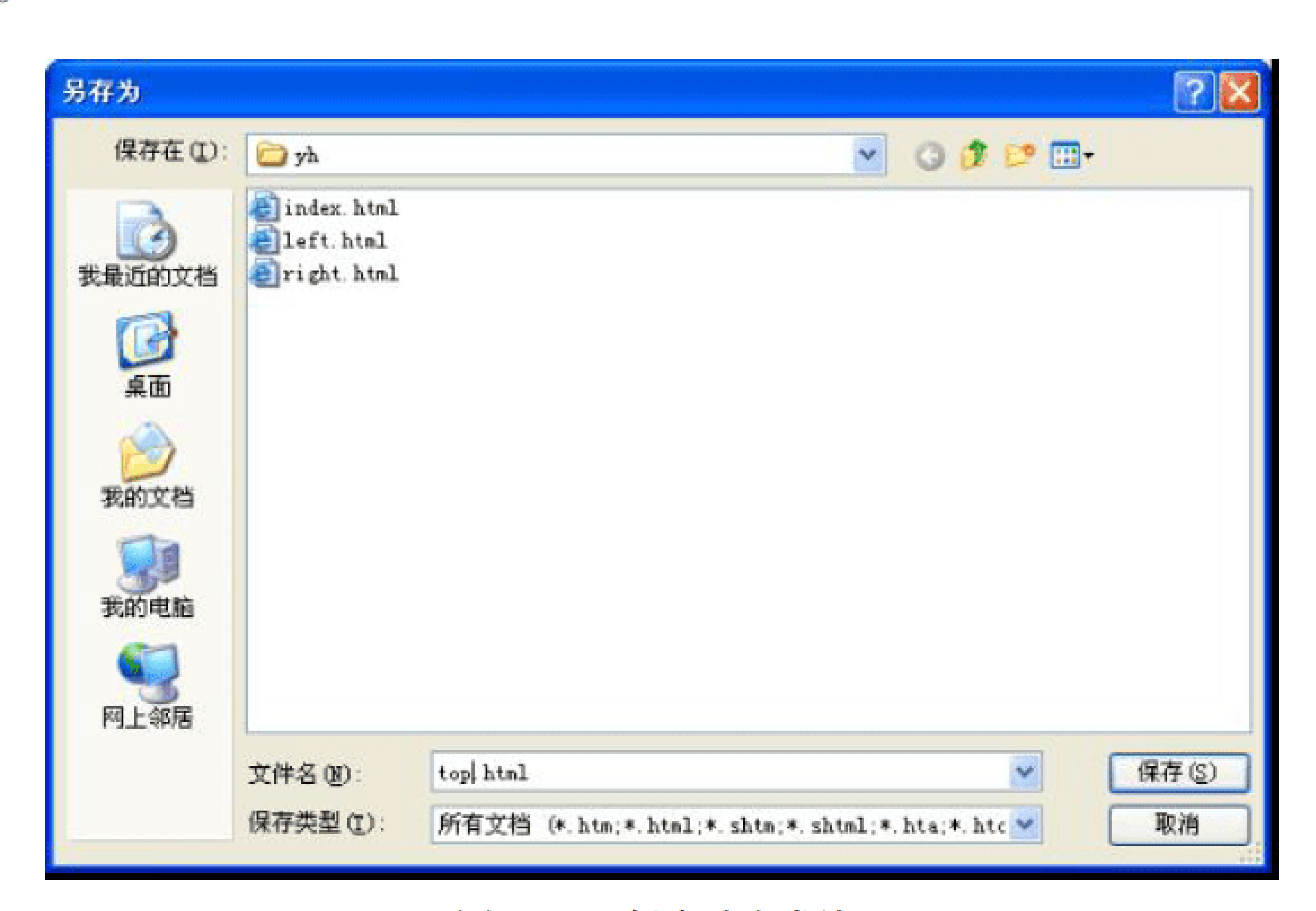


图 6-12　保存头部框架

步骤 05　单击【保存】按钮，至此整个框架保存完毕。

4. 选择框架和框架集

在【框架】面板中选择框架或框架集的具体操作步骤如下：

步骤 01　选择菜单中的【窗口】|【框架】命令，打开【框架】面板，在面板中单击需要选择的框架，框架的边界就会被虚线包围，如图 6-13 所示。

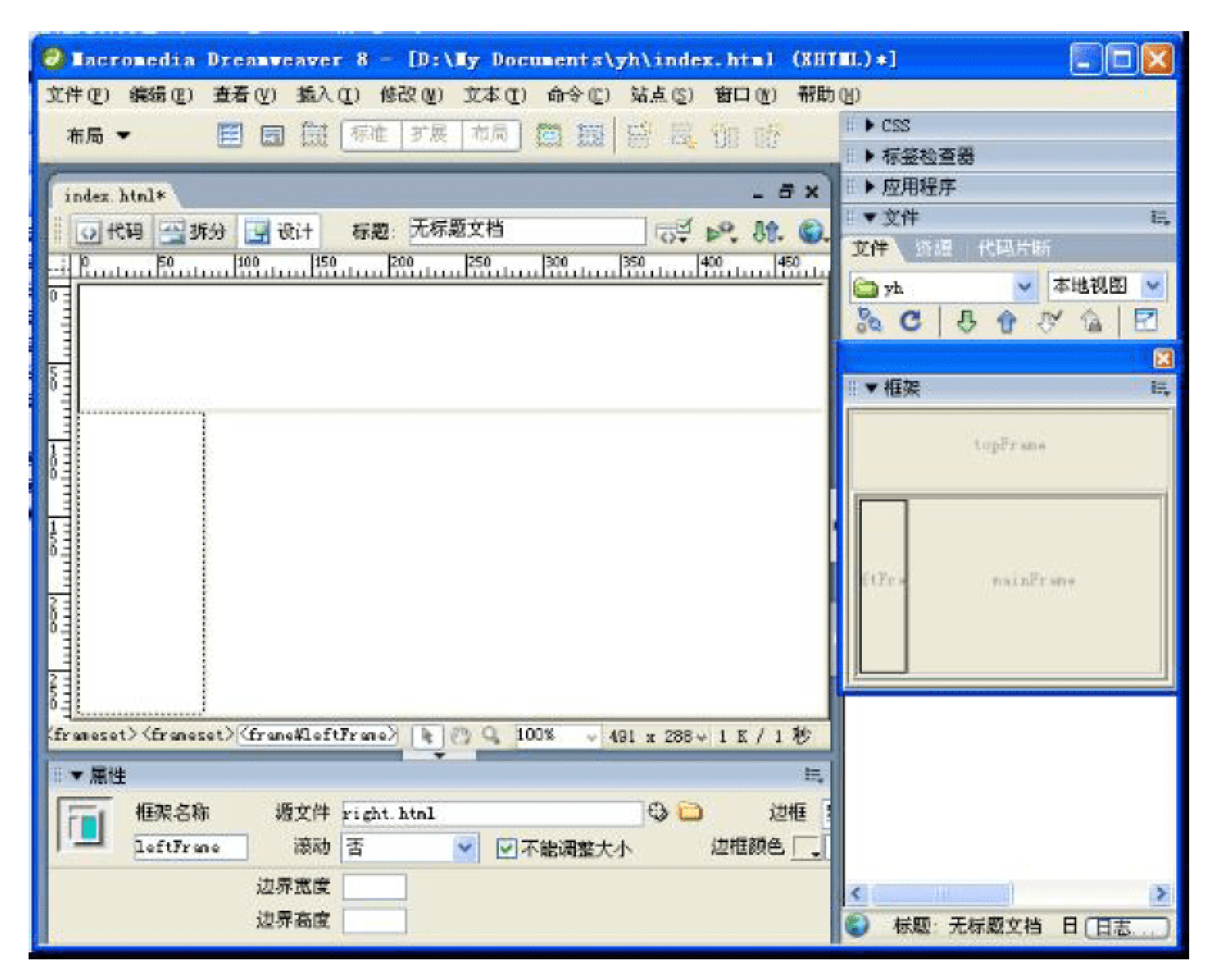


图 6-13　选择需要的框架

步骤 02　在【框架】面板中单击框架集的边框，也可以选择整个框架集，框架集的内侧出现虚线，如图 6-14 所示。

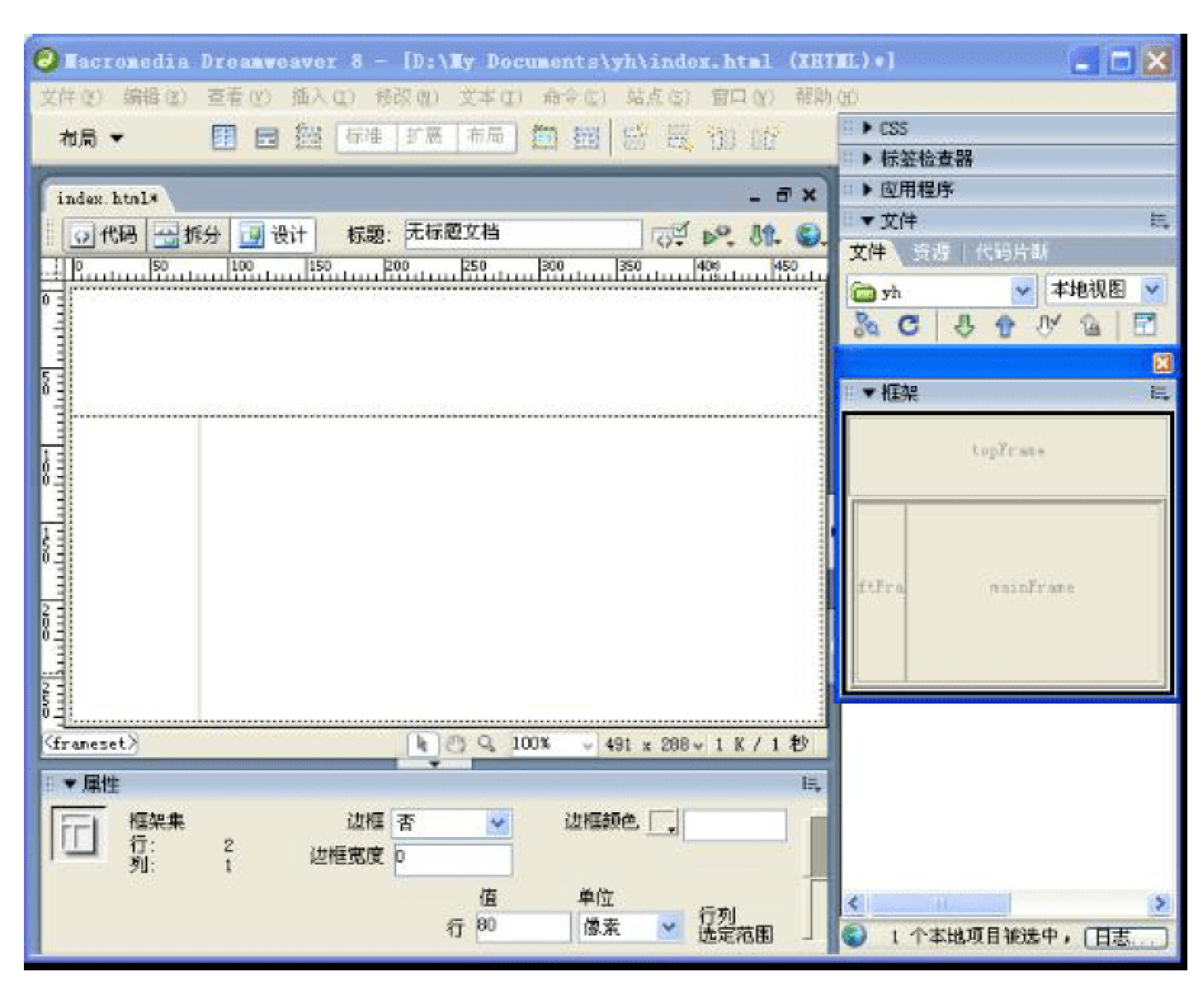


图 6-14　选择整个框架

在文档窗口中选择框架或框架集的具体操作步骤如下：

步骤01 将光标放置在要选择的框架中，按住 Alt 键单击鼠标左键，框架边框内侧出现虚线，即可选择该框架。

步骤02 将光标放置在任意框架中，按住 Alt 键，当鼠标指针靠近框架边线时，出现上下或左右箭头时，单击鼠标左键，框架集内侧出现阴影，即可选中整个框架集，如图 6-15 所示。

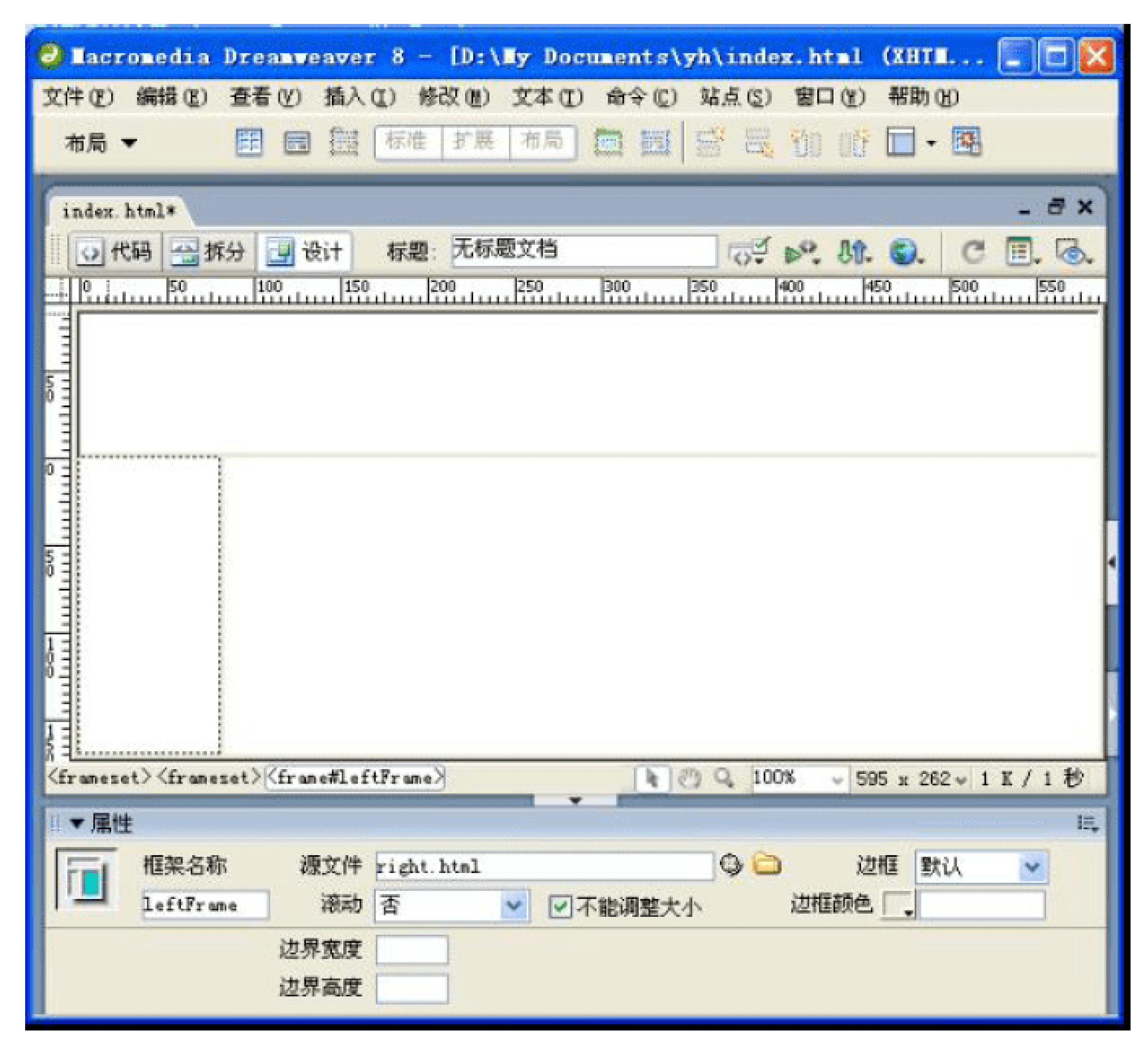


图 6-15　选中整个框架集

5. 设置框架和框架集属性

（1）设置框架属性

使用框架的【属性】面板可以查看和设置大多数的框架属性。选中一个框架，选择菜单中的【窗口】|【属性】命令，打开【属性】面板中显示选中的框架属性。框架的【属性】面板参数如下，如图 6-16 所示。

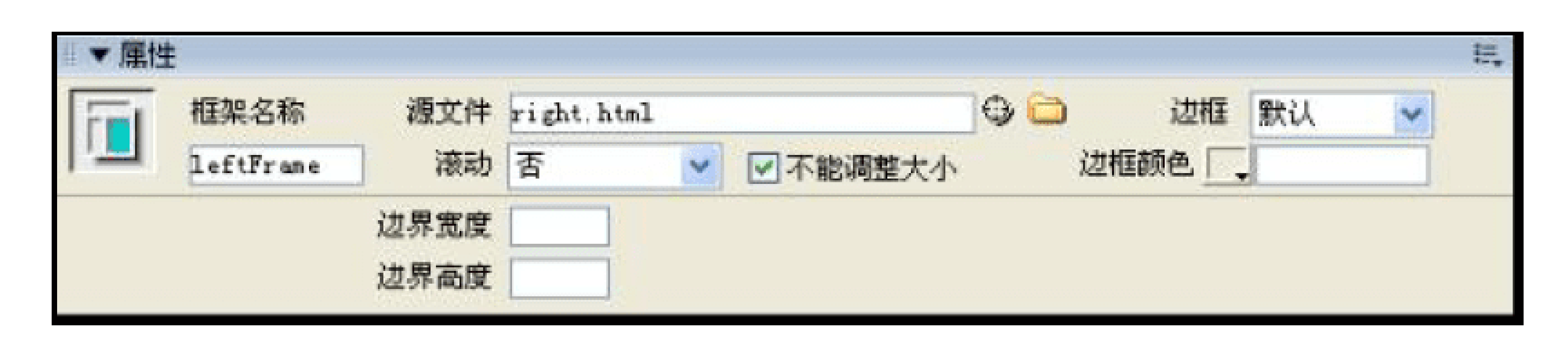


图 6-16　框架属性

框架【属性】面板中各选项说明如下。

- 框架名称：用于为当前框架命名。
- 源文件：确定框架的源文档，可以直接输入名字，或单击文本框右边的【打开】按

钮，可查找并选取文件。

- 滚动：当框架内的内容显示不完整的时候是否出现滚动条，包括是、否、启动和默认4个选项。
- 不能调整大小：限定框架尺寸，防止用户拖动框架边框。
- 边界宽度：设置以像素为单位的框架边框和内容之间的左右边距。
- 边界高度：设置以像素为单位的框架边框和内容之间的上下边距。
- 边框：用来控制当前框架边框，包括是、否和默认3个选项。
- 边框颜色：设置与当前框架相邻的所有框架的边框颜色。

（2）设置框架集属性

选中框架集，打开【属性】面板，在【属性】面板中设置框架集的属性。框架集的【属性】面板参数如下，如图6-17所示。



图6-17 设置框架集属性

框架集【属性】面板中各选项说明如下。

- 边框：设置是否有边框，包括是、否、默认3个选项。选择默认选项，将由浏览器端的设置来决定。
- 边框宽度：设置整个框架集的边框宽度，以像素为单位。
- 边框颜色：用来设置整个框架集的边框颜色。
- 行或列：用来设置整个框架集的行高或列宽的值。
- 单位：行、列尺寸的单位，包括像素、百分比、相对3个选项。

（3）技巧提升

①如何隐藏滚动条

在框架的【属性】面板中设置【边框】和【滚动】都为否，框架的边框是隐藏的，勾选【不能调整大小】复选框，即可隐藏滚动条。

②如何调整框架边框的粗细

选择整个框架集，打开框架集的【属性】面板，将【边框宽度】设置为0，则不显示边框，若想要调整边框粗细，则设置【边框宽度】即可。

③如何准确地选择所需的框架

在制作框架网页时，往往不能够准确地选择某个框架集，对制作框架网页来说非常不方便，准确地选择所需的框架的具体操作方法如下：选择菜单中的【窗口】|【框架】命令，打

开【框架】面板，在面板中单击想要选择的框架，框架的边界就会被虚线包围。

④如何调整框架大小

选中要调整大小的框架集，选择菜单中的【窗口】|【属性】命令，打开【属性】面板，在面板的【行列选择范围】显示框中单击左侧标签或顶部标签分别选定要调整框架所在的行或列，然后在【值】文本框中输入数值，在后面的下拉列表中选择【值】的单位。

⑤如何实现单击一个链接即可改变两个框架中的内容，其具体操作步骤如下：

步骤01 在左侧框架中选中图像，在【属性】面板中选择【矩形热点】工具，在文字上面绘制热点。

步骤02 选中热点，在【属性】面板中的【链接】文本框中输入 right1.html，在【目标】下拉列表中选择 mainFrame。

步骤03 选中热点，选择菜单中的【窗口】|【行为】命令，打开【行为】面板，在面板中单击添加按钮，在弹出的菜单中选择【转到 URL】选项。

步骤04 打开【转到 URL】对话框，在对话框中的【打开在】列表框中选择【框架“topFrome”】，在 URL 文本框中输入 topl.html。

步骤05 单击【确定】按钮，添加行为。

步骤06 保存文档，按 F12 功能键在浏览器中预览效果。

6. 怎样使框架集在不同的浏览器中正常显示

在为以百分比或者相对值指定大小的框架分配空间之前，先为以像素为单位指定大小的框架分配空间。设置框架大小最常用的方法是将左侧框架设置为固定像素宽度，将右侧框架大小设置为相对大小，这样在分配像素宽度后右侧框架就能够伸展，以占据所有的剩余空间。

7. 如何防止网页内容被放在其他网站的框架中

步骤01 切换到拆分视图，在<head>与<／head>之间相应的位置输入以下代码。

```
<SCRIPT LANGUAGE=JAVASCRIPT>
<!--
if (top.location !== self.location) {
top.location=self.location;
 }
</SCRIPT>
```

步骤02 在<body>语句中输入代码 onload=“checkframe();”。

步骤03 保存文档，按<F12>键在浏览器中预览效果。

设置框架大小时，3 种单位优先级的顺序从高到低依次为像素、百分比、相对。如果将某个框架大小以像素为单位进行了设置，然后又用其他单位设置了其框架的大小，则 Dreamweaver 会首先保留当前以像素为单位所设置的框架大小。

8. 如何进行框架的嵌套

步骤01 将光标放置在需要进行嵌套操作的父框架中，然后选择菜单中的【修改】|【框架页】命令，在弹出的子菜单中选择需要的操作，就可以在父框架内嵌套所选格式的框架集。

步骤02 将光标放置在需要进行嵌套操作的父框架中，然后单击【布局】插入栏中的预定义框架集，即会在原父框架内套进一个所选的框架集。

6.6 Flash

ActionScript 代码控制是全 Flash 网站实现的关键，尤其是 Flash MX 新增了很多功能强大的命令，关于这部分，大家可以参看其他的资料，这里只介绍部分制作全 Flash 网站需要使用的比较重要的 ActionScript 代码。

1. 使用 ActionScript 代码

```
loadMovieNum ("url",level[, variables])
loadMovie ("url",level/target[, variables])
```

功能说明：在播放原来加载的影片的同时将 SWF 或 JPEG 文件加载进来。

参数说明：

- url 要加载的 SWF 或 JPEG 文件的绝对或相对 URL，不能包含文件夹或磁盘驱动器说明。
- level 把 SWF 文件以层的形式载入到 Movie 中，若载入 0 层，则载入的 SWF 文件将取代当前播放的 Movie，2 层高于 1 层。
- target 可用路径拾取器取得并替换目标 MC，载入的电影将拥有目标 MC 的位置、大小和旋转角度等属性。
- variables 为可选参数，指定发送变量所使用的 HTTP 方法（GET/POST），如果没有则省略此参数。
- Level 为层次。Flash 允许同时运行多个 SWF 文件，Flash 一旦载入一个 SWF 文件，则占据了一个“层次”，系统默认的是_Flash0 或_Level0，之后的 Movie 则按顺序放在 level0~level16000 里。第一个载入的 SWF 文件为_Flash0 或_Level0，第二个如果加载到第一层时的称为_Flash1 或_Level1，依此类推。注意，前提是前面载入的文件没有退出，否则退出第一个 SWF 文件，第一个文件也从内存中退出。

注意，如果你将外部的 Movie 加载到 Flash0 层或者 Level0 里，那么，原始的 Movie 就会被暂时取代，要再用时还得重新 Load 一次，也就是说，一个 Level 在一个时间里只能有一个

Movie 存在。在使用 LoadMovie 和 UnLoadMovie 时必须特别注意 Level 之间的关系，否则，当你希望在一个时间里只播放一个 Movie 时，就会出现不必要的麻烦。

```
unloadMovieNum (level)
unloadMovie[Num] (level/"target")
```

功能说明：从 Flash Player 中删除已加载的影片。

参数说明：同上

```
loadVariables  ("url" ,level/"target" [, variables])
```

功能说明：从外部文件中（例如文本文件，或由 CGI 脚本、Active Server Page（ASP）、PHP 或 Perl 脚本生成的文本）读取数据，并设置 Flash Player 级别或目标影片剪辑中变量的值。

参数说明：

- url 变量所处位置的绝对或相对 URL。
- level 指定 Flash Player 中接收这些变量的级别的整数。
- target 指向接收所加载变量的影片剪辑的目标路径。
- variables 可选参数，指定发送变量所使用的 HTTP 方法（GET/POST），如果没有则省略此参数。

```
gotoAndPlay (scene, frame)
```

功能说明：转到指定场景中指定的帧并从该帧开始播放。如果未指定场景，则播放时将转到当前场景中的指定帧。

参数说明：

- scene 转到的场景的名称。
- frame 转到的帧的编号或标签。

2. Loading 的制作

考虑到网络传输的速度，如果 index.swf 文件比较大，在它被完全导入以前，设计一个 Loading 引导访问者耐心等待是非常有必要的。同时好的 Loading 在某些时候还可以为网站起一定的铺垫作用。

一般的做法是先将 Loading 做成一个 MC，在场景的最后位置设置标签（如<end>），通过 ifFrameLoaded 来判断是否已经下载完毕，如果已经下载完毕则通过 gotoAndPlay 控制整个 Flash 的播放。

以一个 Loading 文件为例，在场景里加入插入 MC，

```
ifFrameLoaded  ("end" )  {
```

```
gotoAndPlay ("开始播放的地方");
}
```

3. 文本导入

在制作全 Flash 网站的过程中经常遇到一定量的文字内容需要体现，文本的内容表现与上面介绍的流程是一样的，不同的地方体现在最后的表现效果和处理手法上。

（1）文本图形法

如果文本内容不多，有希望将文本内容做得比较有动态效果的，可以采用此法。将需要的文本做成若干个 Flash 的元件，在相应的位置安排好。文本图形法的文件载入与上面介绍的处理手法比较类似，原理都差不多。具体动态效果就有待大家自己去考虑，这里就不多介绍。

（2）直接导入法

文本导入法可以将独立的.txt 文本文件，通过 loadVariables 导入到 Flash 文件内，修改时只需要修改 .txt 文本内容就可以实现 Flash 相关文件的修改，非常方便。

在文本框属性中设置 Var 变量名（注意这个变量名）。

为文本框所在的帧添加 ActionScript 代码：

```
loadVariables ("变量名.txt", "");
```

编写一个纯文本文件.txt（文件名随意），文本开头为“变量名=”，“＝”后面写上正式的文本内容。

6.7 实例制作剖析

一般的制作流程为：网站结构规划→Flash 场景规划→素材准备→分别制作→整体整合。

1. 网站结构规划

网站结构规划的例子，如图 6-18 所示。网站主要结构包括：

网站栏目：News、About、E-mail、Gallery、Cartoon、Animation。

子栏目：Feiying is、Member、Relationship、Contact Us。

图中，实线部分构成主场景（舞台），每个子栏目在首页里仅保留名称，属性为按钮。

虚线部分内容为次场景（演员），可以将次场景内容做在一个文件内，同时也可以做成若干个独立文件，根据需要导入到主场景（舞台）内。

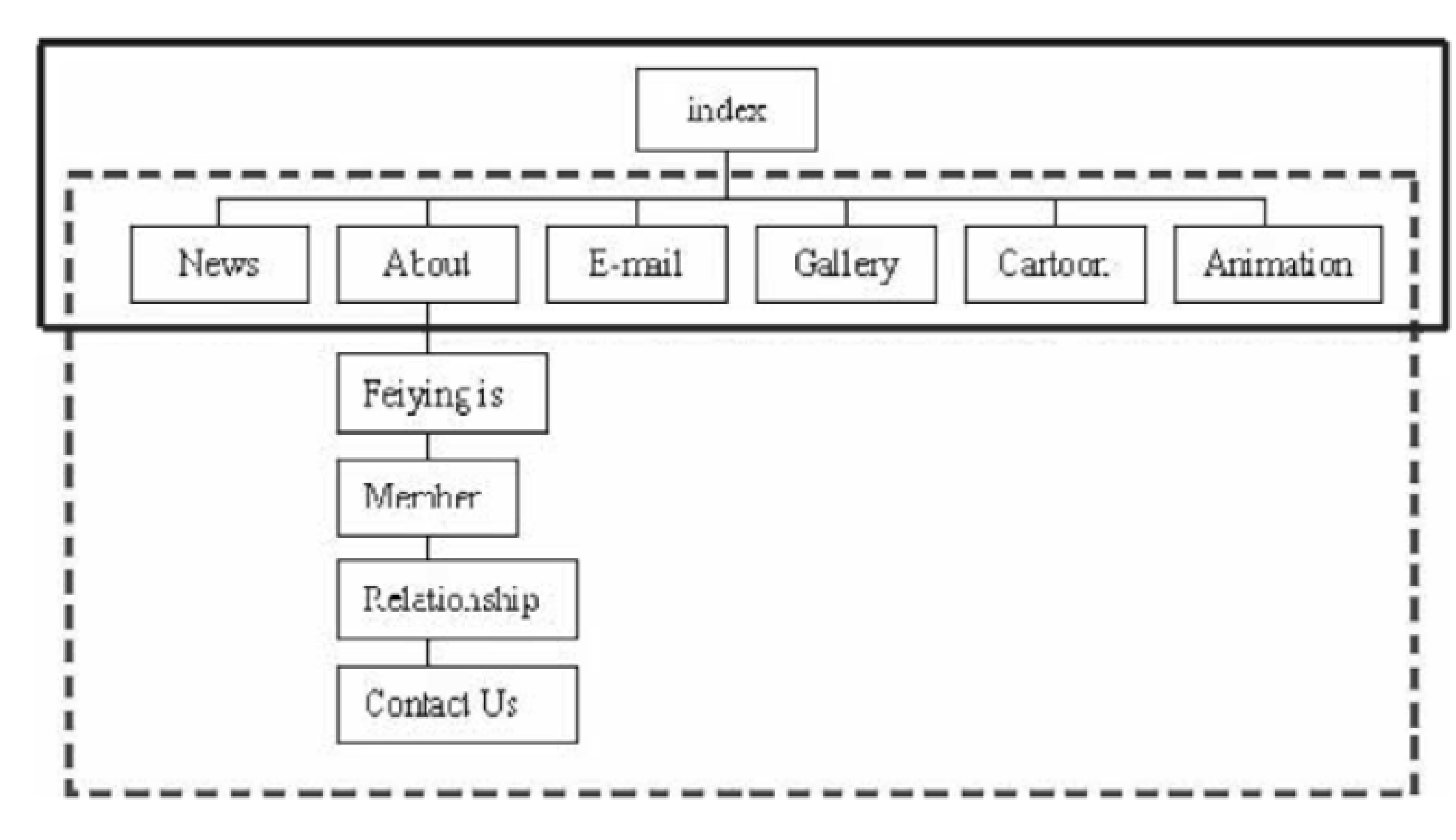


图 6-18　网站结构

2. Flash 场景规划

全网站由主场景、子场景、次子场景……构成。

与制作网站类似，一般会制作一个主场景 index.swf，主要内容包括长宽比例、背景、栏目导航按钮、网站名称等“首页”信息。最后发布成一个 html 文件，或者自己做一个 html 页面，内容就是一个表格，里面写上 index.swf 的嵌入代码即可。

主场景安排，如图 6-19 所示。

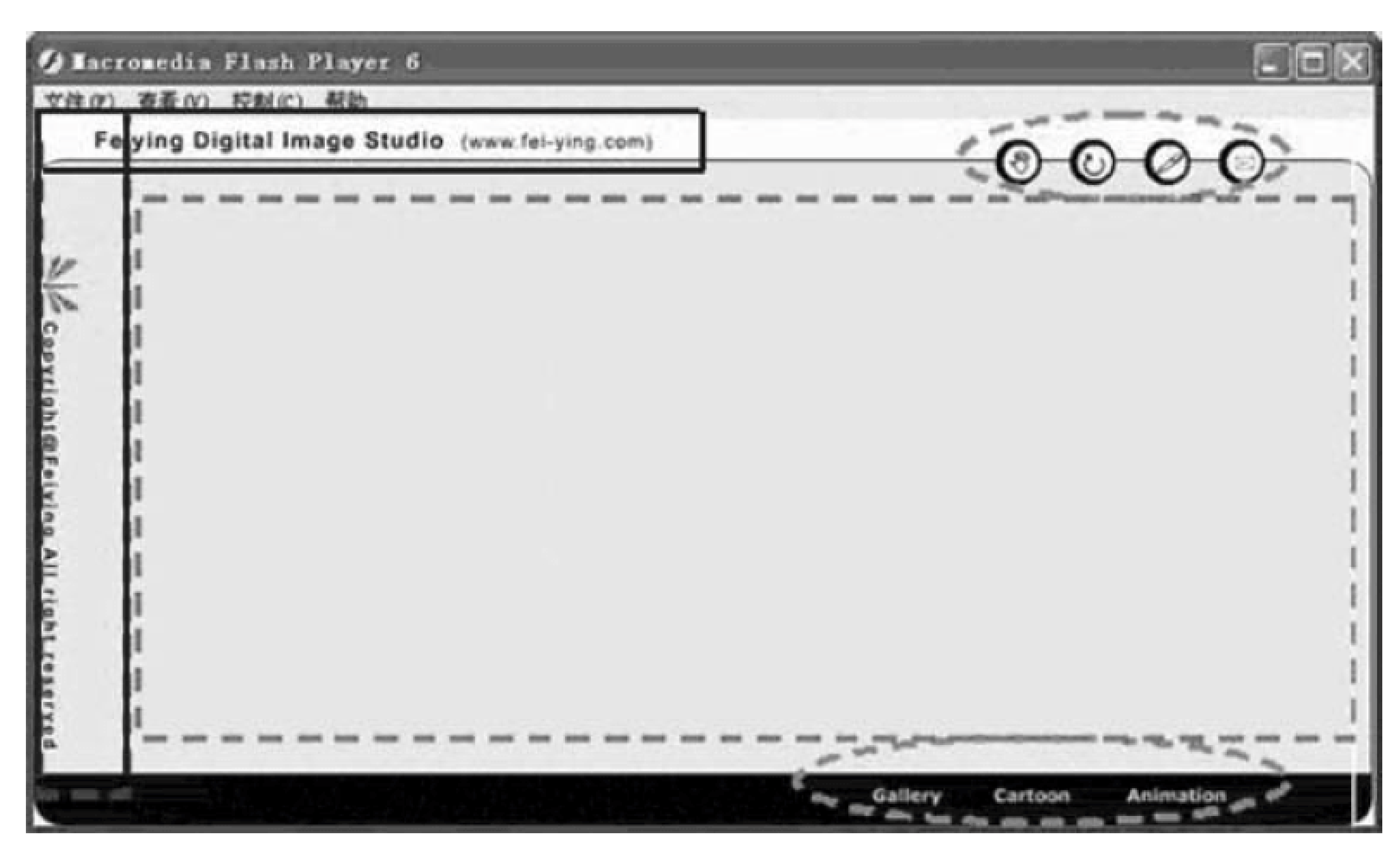


图 6-19　主场景安排

图中，实线部分为网站名称、版权等固定信息区，通常所在位置为 Flash 动画的边缘位置。

椭圆虚线部分为网站栏目导航按钮，通常也是固定在某个区域。按钮可以根据需要做成静态或动态效果，甚至可以做成一个包含 MC 变化的 Button。

虚线部分为主场景导入子文件的演示区域。

在子文件的装载方面主要用到 LoadMovieNum、UnloadMovieNum 两个控制函数，请参看前文的 ActionScript 代码控制。

这里以子栏目 Cartoon 的制作为例。主场景文件 index 中有一个按钮 Cartoon，当我们点击 Cartoon 按钮时希望导入 cartoon 文件夹下的 200208.swf 文件。所以我们在场景内选择 Cartoon 按钮，添加 Action 代码：

```
on (release) {
loadMovieNum("cartoon/200208.swf", 1);
unloadMovieNum (2);
}
```

3. 素材准备

现在确定 Cartoon 子栏目需要导入的文件 200208.swf，该文件计划包含 5 个子文件。所以 200208.swf 文件的界面只包含用于导入 5 个独立子文件的 5 个图形按钮和一个标题。

从图 6-20 可以看到，200208.swf 文件包含 5 个属性为 Button 的小图标，分别为 Bu_2_frog01~Bu_2_frog05。点击它们则分别导入相应文件 200208_frog01.swf~200208_frog05.swf 文件。



图 6-20　需要导入的素材

在场景内选择 Bu_2_frog01，为这个按钮添加 ActionScript：

```
on (release) {
loadMovieNum("cartoon/200208_frog01.swf", 2);
}
```

点击 Bu_2_frog02，为这个按钮添加 ActionScript：

```
on (release) {
loadMovieNum("cartoon/200208_frog02.swf", 2);
}
……
```

依次将 5 个 Button 分别设置好相对应的 Action 以便调用相应的文件。

注意，这里我们设置 level 为 2，是为了保留并区别主场景 1 而设置的导入的层次数，如果需要导入下一级的层数，则层数增加为 3，依次类推。

4. 分别制作

这里的二级次场景是与上级关联的内容，是本例三级结构中的最后一级。该级主要为全 Flash 网站具体内容部分，可以是详细的图片、文字、动画内容。这里需要连接具体的图片为内容，但同样需要做成与主场景比例同等的 .swf 文件，如图 6-21 和图 6-22 所示。



图 6-21　200208_frog01



图 6-22　200208_frog02

该场景是最底层场景，为主体内容显示部分，具体动画可以根据需要做得更加深入。注意要在场景最后一帧处加入停止 ActionScript 代码：stop（）；这样可以停止场景动画的循环动作。

完整导入图片到主场景内的效果，如图 6-23 所示。



图 6-23　完整导入图片后的效果

5. 整合整体

查看本例的 About 子栏目，在文件 index.fla 里设置 About 按钮的 Action：

```
on （release） {
loadMovieNum（"aboutus.swf", 1）;
unloadMovieNum （2）;
}
```

在 aboutus.fla 文件中做好显示文本的文本框，文本框属性设置为多行（Multiline），变量名为 Var:aboutus（注意这个变量名）。

为文本框所在的帧加 ActionScript 代码：loadVariables（"aboutus.txt", ""）；在 aboutus.swf 文件所属目录下编写一个纯文本文件 about.txt，文本开头为“aboutus=”，“＝”后面写上正式的文本内容，如图 6-24 所示。



图 6-24　文本导入

将文本文件完整导入到主场景内的效果，如图 6-25 所示。



图 6-25　完整导入文本后的效果

此时需要注意以下几点。

（1）注意所有子文件的长、宽属性。全 Flash 网站从画面层次来看，非常类似 psoto 的层结构，可以把每个子场景看作一个层文件，子文件是在背景的长宽范围内出现。为了方便定位，可以让子文件与主场景保持统一的长宽比例，这样非常便于版面安排。否则就必须用 setProperty 语句小心控制它们的位置。

（2）发布文件时注意将 HTML 选项发布为透明模式。需要将每个子文件发布为透明模式的原因是不能让子文件带有背景底色，由于子文件的长宽比例与主场景基本是一致的，如果子文件带有底色，就会遮盖主场景的内容。

设置方法：在发布设置里勾选 HTML 选项，在 HTML 选项卡里选择 windows mode：Transparent Windowless，如图 6-26 所示。

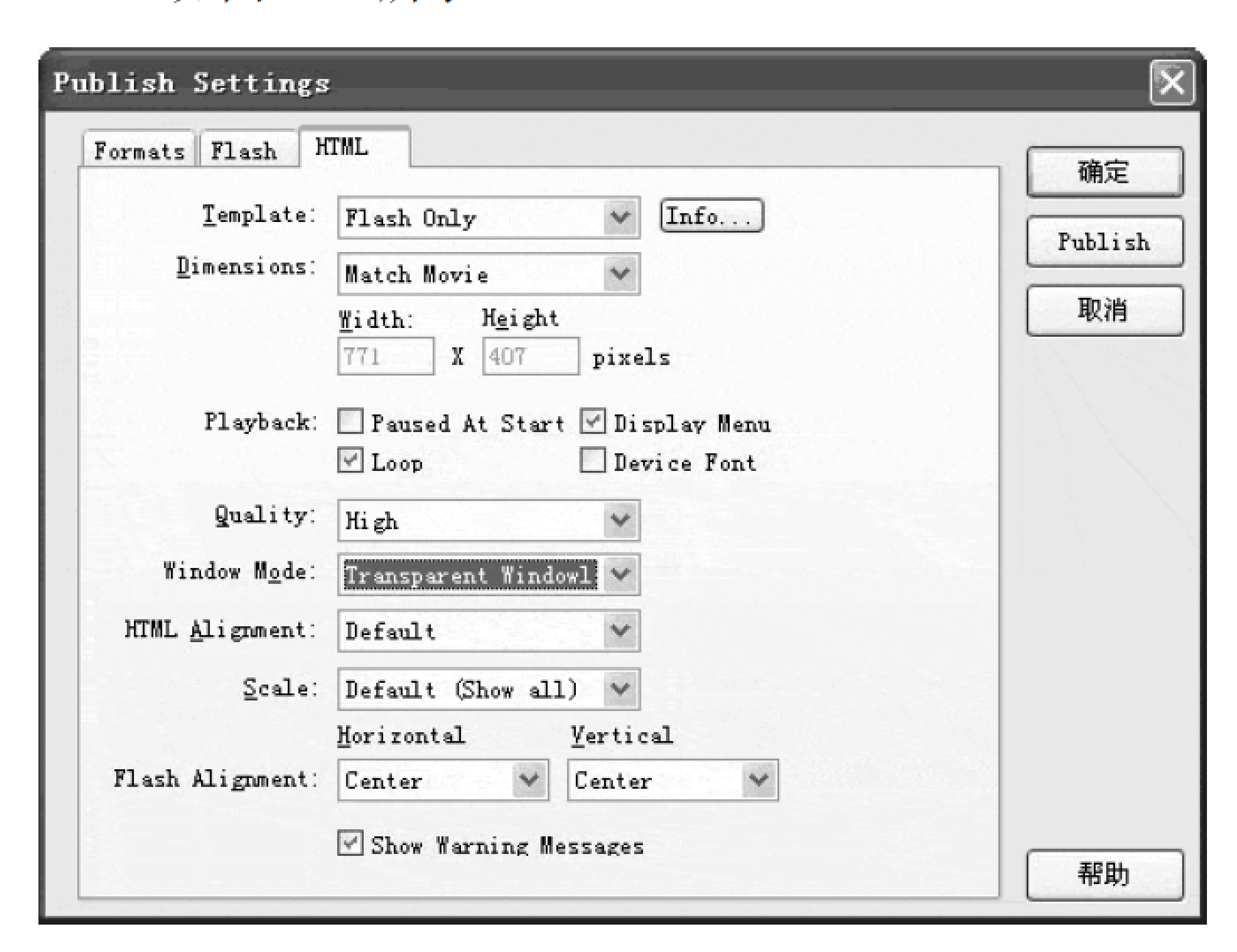


图 6-26　发布设置

（3）使用文本导入时，注意文本文件开头的内容必须是与文本框属性中 Var 定义名相

同的字符串=“正文”。另外需要导入文本的 .swf 文件与被导入的 .txt 文本文件最好在同一目录内。

（4）注意仔细检查文件之间的调用是否正确，避免出现“死链接”。

6.8 小结

页面结构的构建是页面最基本的工作，页面结构经过调整后可以有效地提高页面的相关性及用户体验。另外，在规划页面时，要尽量避免使用对搜索引擎并不友好的技术，如隐藏层、Flash 及框架等。

第 7 章　网站结构优化

目前已经很少有网站是由单独的页面组成的，一般网站都会有几个目录（频道）、无数网页。对搜索引擎优化而言，当涉及多个尤其是成千上万个网站页面时，往往就需要有清晰的网站结构，以确保搜索引擎和用户可以方便地进行访问。而网站结构的规划，对各级关键词优化力度的分配、网页权重的传递都起着非常重要的作用。

7.1 网站物理结构和逻辑结构

网站结构就是指在网站中页面与页面之间的层次关系，按性质可以分为逻辑结构和物理结构。合理的网站结构是能正确表达网站的基本内容及内容之间的层次关系，使得用户、搜索引擎在浏览网站时可以方便地获取信息，不至于迷失。

对搜素引擎优化而言，优秀的网站结构优化应该包括网站物理结构优化和逻辑结构优化两个方面。

很多刚刚接触 SEO 的新手，往往不清楚物理结构和逻辑结构的异同，进而造成自己的网站结构规划很乱，不但搜索引擎不容易快速建立起网站的整体结构体系，访问者更是很难找到自己想要的资料，这无疑对 SEO 起着非常不好的作用。

下面简单介绍网站物理结构和逻辑结构的实际例子，以便读者快速理解、熟悉这两种典型结构。

7.1.1 网站物理结构

网站物理结构指的是网站目录及所包含文件所存储的真实位置所表现出来的结构，物理结构一般包含两种不同的表现形式：扁平式物理结构和树形物理结构。

1. 扁平式物理结构

对于小型网站来说，所有网页都存放在网站根目录下，这种结构就是扁平式物理结构。

采用扁平式物理结构的网站，页面 URL 地址一般如下：

- http://www.xxx.com/index.html

- http://www.xxx.com/wenzhang1.html
- http://www.xxx.com/wenzhang2.html
- http://www.xxx.com/wenzhang3.html

这种扁平式物理结构对搜索引擎而言是最为理想的，因为只要一次访问即可遍历所有页面。但是，如果网站页面比较多，则太多的网页文件都放在根目录下，这样查找、维护起来就显得相当麻烦，所以，扁平式物理结构一般适用于只有少量页面的小型、微型站点。

2. 树形物理结构

对规模大一些的网站，往往需要两到三层甚至更多层级子目录才能保证网页的正常存储，这种多层级目录也叫做树形物理结构，即根目录下再细分成多个频道或目录，然后在每一个目录下再存储属于这个目录的终极内容网页。

采用树形物理结构的网站，页面地址一般如下：

- http://www.xxx.com/index.html
- http://www.xxx.com/lanmu1/
- http://www.xxx.com/lanmu2/
- http://www.xxx.com/lanmu3/
- http://www.xxx.com/lanmu1/wenzhang1.html
- http://www.xxx.com/lanmu1/wenzhang2.html
- http://www.xxx.com/lanmu2/wenzhang1.html
- http://www.xxx.com/lanmu2/wenzhang2.html
- http://www.xxx.com/lanmu3/wenzhang1.html
- http://www.xxx.com/lanmu3/wenzhang2.html

采用树形物理结构的好处是维护容易，但是搜索引擎的抓取将会显得相对困难。目前互联网上的网站，因为内容普遍比较丰富，所以大多都是采用树形物理结构。

7.1.2 网站逻辑结构

网站的逻辑结构也叫做链接结构，主要是指由网页内部链接所形成的逻辑结构。逻辑结构和物理结构的区别在于，逻辑结构由网站页面的相互链接关系决定，而物理结构则由网站页面的物理存放位置决定，如图 7-1 所示。

在网站的逻辑结构中，一般采用“链接深度”来描述页面之间的逻辑关系。“链接深度”指从源页面到达目标页面所经过的路径数量，如某网站的网页 A 中，存在一个指向目标页面 B 的链接，则从页面 A 到页面 B 的链接深度就是 1。

与物理结构类似，网站的逻辑结构同样可以分为扁平式和树形两种。

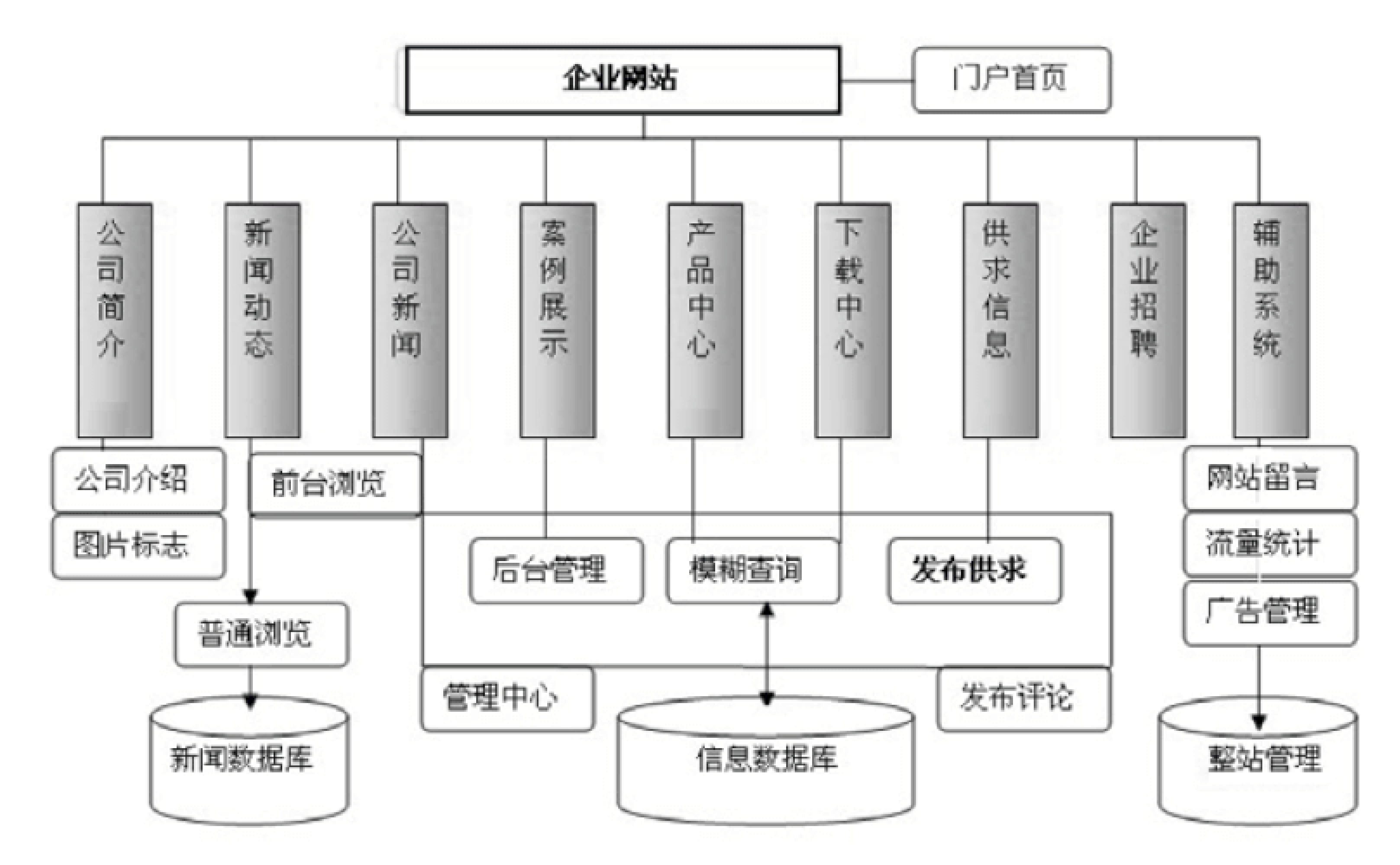


图 7-1 网站的逻辑结构图

1. 扁平式逻辑结构

扁平式逻辑结构的网站，实际上就是网站中任意两个页面之间都可以相互链接，也就是说，网站中任意一个页面都包含其他所有页面的链接，网页之间的链接深度都是 1。目前的网络上，很少有单纯采用扁平式逻辑结构作为整站结构的网站。

2. 树形逻辑结构

是指用分类、频道等页面，对同类属性的页面进行链接地址组织的网站结构。在树形逻辑结构网站中，链接深度大多大于 1。

7.2 网站结构对 SEO 的影响

网站结构对网站的搜索引擎友好度、用户体验都有着非常重要的影响。简单概括起来分以下 3 点。

（1）网站结构可以影响网站内部页面的重要性

网站结构在决定网站内部页面的重要性方面，起着非常关键的作用，如合理的内部链接策略就可以对重要页面进行突出、推荐等操作。

（2）网站结构是衡量网站用户体验好坏的重要指标

清晰的网站结构可以帮助用户快速获取所需的信息，如图 7-2 所示。相反，如果网站结构混乱，用户在访问时就如同走进迷宫，很难找到自己想要的资料，只能离开。



图 7-2　新浪首页清晰的导航

（3）网站结构会直接影响到搜索引擎对页面的收录

一个合理的网站结构可以引导搜索引擎抓取到更多、更有价值的网页。相反，网站结构混乱，往往就会造成搜索引擎陷入死循环、抓取不到页面等问题。

7.3　网站结构优化的基本方法

网站结构优化就是对网站页面的存储方式（物理结构）、内部链接关系（逻辑结构）进行合理的调整，以减少网站的目录层级，优化普通页面与重要页面之间的链接深度，增加重要页面的链接入口，最后达到所有页面都能被搜索引擎方便地收录，并且合理分配网站内部页面权重的目的。

物理结构优化主要是减少网站页面的目录存储深度，一般以 URL 的目录层次作为标准。

一般说来，网站建设者在优化物理结构时，不提倡将所有文件都存放在根目录下，而应该按栏目内容建立子目录，并且尽量使用意义明确的目录名称，比如 image、css、js、post、bbs 等，但是注意目录名称不宜太过冗长，URL 应该越短越好。

目录和文件命名可以适当使用关键词，如果是关键词组，需要用分隔符分开。现在很多 CMS 直接提供以关键词拼音的方式建立文件夹，如图 7-3 所示。

图 7-3　CMS 中的按拼音保存文件夹

虽然建议在优化物理目录时，创建一定的子目录来进行页面的分类和整理，但是网站的目录层级一般建议不要超过 3 层。如很多 CMS 默认生成的页面存储结构如下：

- Http://www.xxx.com/html/seo/seo.html
- Http://www.xxx.com/html/sem/sem.html
- Http://www.xxx.com/html/chuangye/chuangye.html

在物理结构优化中，应该去掉上述目录中的“html”，以产生以下的 URL：

- Http://www.xxx.com/seo/seo.html
- Http://www.xxx.com/sem/sem.html
- Http://www.xxx.com/chuangye/chuangye.html

优化目录的深度其实不单单为 SEO，还有很多其他好处，如较少层级的目录更容易被访问者记忆，更容易被传播，更容易交换链接等有益于用户体验的地方。

优化物理结构的长度其实就是在允许的情况下，对网站目录的改写、缩写或者同义替换。举例来说，如果网站的一个目录是 SEO，那么一般没有必要将 SEO 这个目录改成 sousuoyinqingyouhua，因为搜索引擎懂 SEO 这个词的意思。

7.4 逻辑结构优化的基本方法

逻辑结构优化主要是减少页面之间的链接深度，包括减少普通页面与重要页面之间、重要页面与重要页面之间的链接深度，以及为网站中相对重要的页面增加更多的链接入口等方面。

在相对重要的页面中，如果存放更多的链接指向其他页面，不但可以减少页面之间的链接深度，还可以引导搜索引擎抓取网站中更多的页面，传递权重。

网站首页一般是最重要的，在网站首页中，适当地增加内页的指向链接，有助于搜索引擎抓取内页。在现在的 CMS 系统中，网站首页一般都会产生最新更新的内容链接，这样就能让搜索引擎快速发现网站的最新页面，并且快速收录。

7.4.1 增加重要页面的链接入口

在更多的网站内部页面中增加指向重要页面的链接，可以有效增加网站中相对重要页面的链接入口，从而增加这些页面的链接权重。

在网站的结构优化中，片面地理解物理结构和逻辑结构的优化都是不行的，应当综合考量。如不管内容页所处的目录层级有多深，只要该页有一个或者大量来自其他网站的外部链接和内部网站的大量内部链接，它同样会被搜索引擎快速发现，并给予较高的权重。

7.4.2 将需要重点优化的几个词作为网站主题

在了解该网站的主要关键词体系的基础上，详细介绍与关键词体系相对应的网站结构优化这个过程遵循先易后难的法则，从最容易理解的主关键词、二级关键词结构优化开始，再到最后比较繁琐的三级关键词或者长尾关键词的网站结构规划。

7.5 确定网站主要关键词

以“某宠物网”为例，通过关键词体系的建设，已经得到了网站最重要的几个关键词，也可以称为一级关键词。

“某宠物网”这是网站准备建设的网站品牌，也就是促使更多来访者记忆的网站名字。加深来访者对这个名字的记忆，可以极大地提高网站的品牌力量，让网站拥有更多直接访问的回头客。

“宠物”是网站的主题，也就是说，整个网站的所有页面，从内容上说，都应该和这个主题相关。

在通常情况下，网站的主要关键词不宜设置过多，一般一、两个就足矣。常规的做法是一个内容主题，一个品牌。举例来说，新浪网的主要关键词可以看成是品牌名称“新浪”，以及主题“新闻”。

获得了网站的主要关键词以后，搜索引擎优化者应该将这些主要关键词应用于网站结构优化。

7.5.1 与主要关键词相关的网站结构规划

从整个网站结构规划来看，最重要的主关键词自然需要简练的物理结构和最高权重的逻辑链接指向来进行优化，而在当前的主流 CMS 系统中，这两点都很容易实现。

针对站名（品牌）和内容主题（主关键词），从物理结构规划角度出发，应该将它们放置在最容易被搜索引擎发现的地方、被最多页面链接指向的地方，也就是网站主页，如图 7-4 所示。

从逻辑结构规划来看，主关键词无疑需要更多的链接支持和权重传递，所以需要将这些词汇融合到网站的每个页面中，如将站名和主题放置在所有页面的标题中。

从链接权重传递的角度来说，整个网站的每个页面，都需要对主题所在的主要页面进行链接指向，一般情况下，也就是所有页面都需要指向首页。

上面的描述看起来比较繁琐，实际上主流的 CMS 系统都很好地兼容了这一点。



图 7-4　新浪首页

CMS 可以很方便地将站名、主题设置成网站名称，然后在任何页面的标题中都能出现这些词汇。

在默认情况下，CMS 的所有页面都会指向网站首页，不管是文字还是图片类型的链接都有。

网站站名和主题的链接规划是最简单，也是最基础的网站结构规划，操作起来也比较简单。下面开始进入稍微复杂一些的二级关键词结构规划。

7.5.2　列二级关键词目录（网站频道规划）

网站的主要关键词及网站品牌（通常是站名）的确定和规划比较简单，在成熟的 CMS 系统的基础上很容易做好。与此类似，网站的二级关键词和相对应的网站结构规划也比较简单，归纳起来分为列出关键词目录和规划网站结构两步。

7.5.3　确定网站二级关键词目录

一般来讲，大型站点的二级关键词根据网站的规模可能有几百个（如新浪），中小型网站的二级关键词不会太多，数量大约有几十个或者 10 个以内。

以某养生网为例，二级关键词有以下几个：

- 中医养生
- 四季养生
- 养生之道
- 养生堂
- 平衡养生说
- 免费养生咨询

这些关键词实际上就是网站后期需要主推的词，从流量角度来讲，也就是比较容易长期、稳定地吸引访问的词汇。

在获得了二级关键词目录以后，下一步就可以规划与之相对的网站结构了。

7.5.4　与二级关键词相对应的网站结构规划

获得了二级关键词以后，要规划网站结构就比较简单了，通常比较实用的方法就是用网站的主要栏目（或者说频道、大分类）来规划、承载这些关键词。

在现有的 CMS 系统中，要将二级关键词作为系统栏目，操作非常简单，下面简单地介绍一下。以国内知名的 CMS 系统织梦 CMS 为例，操作步骤如下：

步骤 01　输入页面地址 http://www.XXXX.com/dede/login.php，再输入账户密码即可进入后台管理首页，如图 7-5 所示。



图 7-5　织梦后台管理

步骤 02　选择左侧快捷菜单中的“网站栏目管理”链接，即可定义栏目，如图 7-6 所示。



图 7-6　选择链接

步骤 03　单击“增加顶级栏目”功能按钮，如图 7-7 所示，即可增加一个栏目的相关操作，分别是“常规选项”、“高级选项”和“栏目内容”。

内容管理系统　频道管理　HTML更新　系统管理　管理主页　网站主页　注销登录

频道管理

网站栏目管理(在IE浏览器中可使用右键菜单进行操作。)　↑创建顶级栏目

选择	ID	名称	记录数	内容模型	管理	排序
	[ID:2]	计算机软件著作权登记	(文档：31)	1	预览 内容 添加 修改 移动 删除	1
	[ID:7]	作品著作权及合同登记	(文档：20)	1	预览 内容 添加 修改 移动 删除	2
	[ID:3]	双软登记	(文档：33)	1	预览 内容 添加 修改 移动 删除	3
	[ID:4]	高新技术企业登记	(文档：36)	1	预览 内容 添加 修改 移动 删除	4
	[ID:6]	联系我们	(文档：1)	1	预览 内容 添加 修改 移动 删除	5
	[ID:11]	三户视觉资讯中心	(文档：53)	1	预览 内容 添加 修改 移动 删除	50
	[ID:9]	常用下载	(文档：4)	1	预览 内容 添加 修改 移动 删除	50
	[ID:10]	图片轮播	(文档：3)	1	预览 内容 添加 修改 移动 删除	50

增加顶级栏目　更新排序　获取JS　更新栏目HTML　更新文档HTML

图 7-7　增加栏目

下面以实际的例子来说明这几个定义项目和网站结构优化的内在联系，并介绍一些栏目关键词的优化细节。

1. “常规选项”与结构规划

在“常规选项”中，可以根据需要创建的二级关键词栏目类型定义相关参数。这一步需要注意的是网站的物理目录结构规划方式，通常栏目需要放置在网站根目录下，也就是能产生 http://www.XXXX.com/lanmu/这样的 URL 物理结构。

另外，栏目的文件夹名称需要根据实际的 SEO 策略进行定制，如二级关键词比较短小，而且主要目标是做中文搜索引擎的 SEO，则通常采用拼音的方式作为栏目名，这对优化有一定的好处。

2. “高级选项”与结构规划

在“高级选项”中，可以定义二级关键词对应栏目的逻辑 URL 方式，比如是否启用二级域名支持。如果网站总体规划相对较大，内容足够充分，则可以采用二级域名的方式规划，如图 7-8 所示。

除了域名方式以外，在“高级选项”中，还可以定义这个栏目下的栏目首页、列表页以及内容页的 URL 结构。在网站物理结构规划中，这些 URL 结构也是比较重要的，通常使用的方法是将列表页和内容页直接生成在栏目文件夹下。

在“高级选项”中，还需要注意的是“SEO 标题”、“关键词”和“栏目描述”这 3 个选项。

栏目管理>>增加栏目

常规选项　高级选项

多站点支持：　⦿ 不启用　○ 启用

说明：　绑名绑定仅需要在顶级栏目设定，下级栏目的目录和域名将相对于顶级栏目。

绑定域名：　（需加 http://，一级或二级域名的根网址）

站点根目录：　⦿ 相对于主站点根目录　○ 绝对路径

封面模板：　{style}/index_article.htm　浏览...

列表模板：　{style}/list_article.htm　浏览...

文章模板：　{style}/article_article.htm　浏览...

单独页面模板：　浏览...

图 7-8　规划二级域名

（1）SEO 标题

SEO 标题存在的意义是为了方便网站建设者对现有的栏目标题进行补充。例如，如果一个栏目名（二级关键词）存在同义词或者近义词，在网站的栏目设置中，通常不建议两个词一起作为网站栏目名，而应该选择一个最常被搜索的词作为栏目名（也可以选择最容易优化的一个词），将另外的同义词、近义词通过 SEO 标题的方式规划。

（2）关键词

与 SEO 标题不同的是，“高级选项”中的“关键词”可以有比较灵活的定义方式。

如果网站管理者觉得自己能够规划出与当前栏目内容相符合、有意义并且准确，但又不会产生关键词堆砌的话，可以自定义“关键词”一项。这样有助于搜索引擎理解栏目内容，当用户搜索相关词汇并返回搜索结果时，可以起到返回相关内容并截取部分内容，同时还有促进用户点击的作用。

相反，如果网站管理者在“关键词”选项中，一味地堆砌相关关键词，放置一些和本栏目不相关的词，则建议站长不要写这一项，留空即可。

（3）栏目描述

与“关键词”不同的是，“栏目描述”从用户体验角度而言，应该编写一段和栏目内容相关的介绍，并在介绍中恰当地加入此栏目下的主要内容，用以提升用户体验。

针对不同的栏目产生不同的栏目内容是织梦的一大特色，这个特色非常适合 SEO 采用，对优化栏目和对应的关键词有非常大的帮助。

3. “栏目内容”与结构规划

所谓“栏目内容”，从优化角度讲，就是为这个栏目页产生一段独特的、有意义的内

容，以方便搜索引擎收录、排名。和传统的栏目列表页收录困难相比，织梦的这个功能不但可以实现栏目下的文章列表，还可以打造有具体内容的栏目列表页。

通过上述的设置，与二级关键词相对应的网站栏目结构规划就完成了。在实际的操作过程中，优化者应该以用户体验为基础，有选择性地规划这些栏目的物理结构和逻辑结构，力争让用户可以方便地找到自己所需内容的同时，兼顾搜索引擎蜘蛛的抓取或爬行。

7.6 建三级关键词表

上文介绍了和一、二级关键词相关的网站主页、站名和网站栏目的规划，因为关键词数量较少，所以都比较简单。下面进入比较繁琐且数量比较多的三级关键词相关网站结构规划。

对中小型网站来说，一级关键词可能只有一、两个，二级关键词可能只有几个、数十个，但是三级关键词就比较多了，通常有几百个。如何将数量较大的三级关键词和网站结构优化结合在一起，这就需要在网站上线之前就有完整的三级关键词表。

7.6.1 如何建设三级关键词表

所谓的三级关键词，是相对于网站主关键词和二级关键词而言的，简单来说，就是略微长尾的需要优化的词。以下是三级关键词的构成和特点。

还是以某养生网为例，三级关键词的典型构成。

- 网站主关键词：养生。
- 网站二级关键词中的一个：中医养生。
- 网站三级关键词中的一个：中医体质养生。

三级关键词的特点可以简单分为以下几点。

- 从长度上说，三级关键词通常要比一、二级关键词略长，但是又不是很长，和单纯的长尾关键词又有所区别。
- 从优化难度上说，三级关键词比一、二级关键词更容易优化。
- 从数量上说，三级关键词比一、二级关键词多出十倍、几十倍。
- 从搜索量上说，三级关键词的搜索量一般小于一、二级关键词，但是也有可能存在搜索量比较多的特殊词汇。
- 从用户体验角度而言，搜索三级关键词的用户往往有更明确的需求，这对电子商务和在线营销来说具有更为精准的意义。

一个网站的三级关键词就好比网站金字塔结构中靠近底部的部分：体积比较大，含金量

比较高，操作难度相对较低。

在网站推出之前就建立好三级关键词表与一、二级关键词不同，三级关键词数量很多，存在各种各样的构成方式，这就需要优化者有一份比较详细的三级关键词表。这个表的建立方法有很多种，如通过各种搜索引擎优化工具进行挖掘，通过竞争网站分析获得，通过搜索引擎官方提供的工具搜集（如谷歌关键词工具、百度推广中的相应功能等）。

将网站的三级关键词整理出来，不管是采用 Excel 或者其他数据库进行管理，还是用文档进行记录，这都是网站的三级关键词表。这个表一方面应该在网站推出之前就有比较丰富、全面的内容，另一方面还应该随着网站上线后的不断优化、新热点的出现而继续扩充，一个简单的 Excel 三级关键词表，如图 7-9 所示。

	A	B	C	D	E
1	关键词	百度收录	月均搜索量	昨日搜索量	百度优化指数
2	中医养生	15900000	976	1064	0.61
3	中医养生网	5990000	720	752	1.2
4	中医养生论文	5530000	0	0	0
5	中医体质养生	6790000	172	204	0.25
6	中医养生学	483000	143	142	2.96
7	中医养生保健	6200000	251	231	0.4
8	中医养生美容	5240000	155	177	0.29
9	中医养生美容院	3690000	0	0	0
10	美容院养生项目	227000	0	0	0
11	中医养生保健加盟	3510000	0	0	0
12	中医养生馆	1860000	151	146	0.31
13	中医养生馆连锁加盟	1530000	0	0	0
14	中医养生项目	18000000	0	0	0
15	中医养生馆加盟	15500000	0	0	0
16	中医养生康复	6120000	0	0	0
17	中医养生知识	5550000	0	0	0
18	中医养生的书	6010000	0	0	0
19	中医养生讲座	1070000	174	161	1.62
20	中医美容	6460000	174	326	0.26
21	中医养生视频	897000	181	173	2.01

图 7-9　三级关键词列表

有了三级关键词表，下面就可以着手在网站上进行与之对应的网站结构规划了。

7.6.2　用分类和专题构建三级关键词体系结构

在当前的主流 CMS 系统中，小分类和专题比较适合用于三级关键词体系的结构规划。分类是三级关键词的主要承载方式。

二级关键词一般用栏目或者频道的方式进行规划，三级关键词就比较适合用小分类、二级栏目的方式进行规划。

1. 用分类构建三级关键词体系结构

同样以织梦 CMS 为例，登录网站后台后，可以在相应的大栏目下建立更为细致的二级栏目。这些二级栏目和三级关键词相对应，根据内容的容量、优化的难易度进行规划。如“四季养生”是一个主要的二级关键词，在网站中以栏目的方式呈现，与之对应的三级关键词有“春季养生”、“夏季养生”、“秋季养生”、“冬季养生”等。从网站物理结构规划

上讲，这些三级关键词放置在二级关键词所在的文件夹下比较合理，如图 7-10 所示。

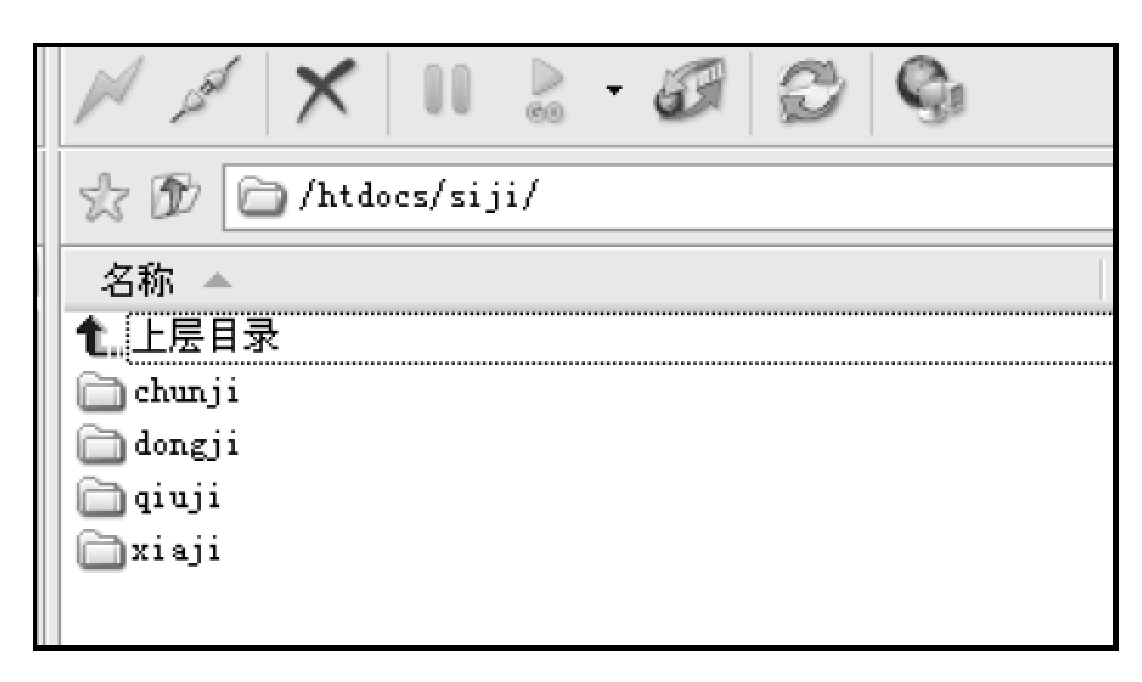


图 7-10 物理位置

这样的三级关键词规划是非常简单的，同时也是优化效果比较好的方式。采用这种方式，可以对二级关键词下的各个栏目进行扩展，按每个主栏目下分出 10～20 个下级栏目为例，一个具有 10 个栏目的网站，主栏目就可以承载 100～200 个三级关键词或者更多。

可能有一些三级关键词从逻辑上讲并不适合放在某个主栏目下，或者是三级关键词分析足够细，使小栏目从数量上说无法承担这么多；另外，从用户体验角度讲，也不想让每个主栏目下面有这么多的小栏目。这些情况都是很常见的，该怎么办呢？有效的解决办法就是利用专题来分担三级关键词的结构规划。

2. 专题构建三级关键词体系结构

所谓专题，在当前主流的 CMS 系统中，就是将任意栏目的文章进行整合，从而归类、聚合出新内容。这个功能在各大 CMS 系统中都有，并不稀奇，但是优化的作用却很大。

当三级关键词比较多，二级栏目无法完成所有的三级关键词结构规划时，站长可以将三级关键词用专题的方式进行分步规划，这样可以大大提升三级关键词的容量；当三级关键词和主栏目存在逻辑归属冲突时，专题可以独立于栏目建立。

当三级关键词并不单单属于某一个主栏目时（也就是说，和多个栏目中的内容存在交集），专题可以很好地聚合这些内容，从而创造出新的内容。

当出现新的热点时，与之对应的三级关键词可能以前并没有得到重视，各栏目内容中也零散地进行了介绍，这时候利用专题的聚合功能就可以一边抓住热点，一边对新关键词进行快速优化。

在当前主流的 CMS 系统中，专题的建立都很简单，下面用实际例子介绍专题建设中的几个注意事项。

步骤 01 进入织梦后台，利用左侧快捷菜单中的“专题管理”功能就可以方便地建立和管理专题，如图 7-11 所示。



图 7-11　专题管理

步骤 02　在“专题管理”中单击“专题列表”链接，即可进行专题的增加操作。通常情况下，一个专题的建立分为两大部分：专题参数设定和文章设定，如图 7-12 所示。

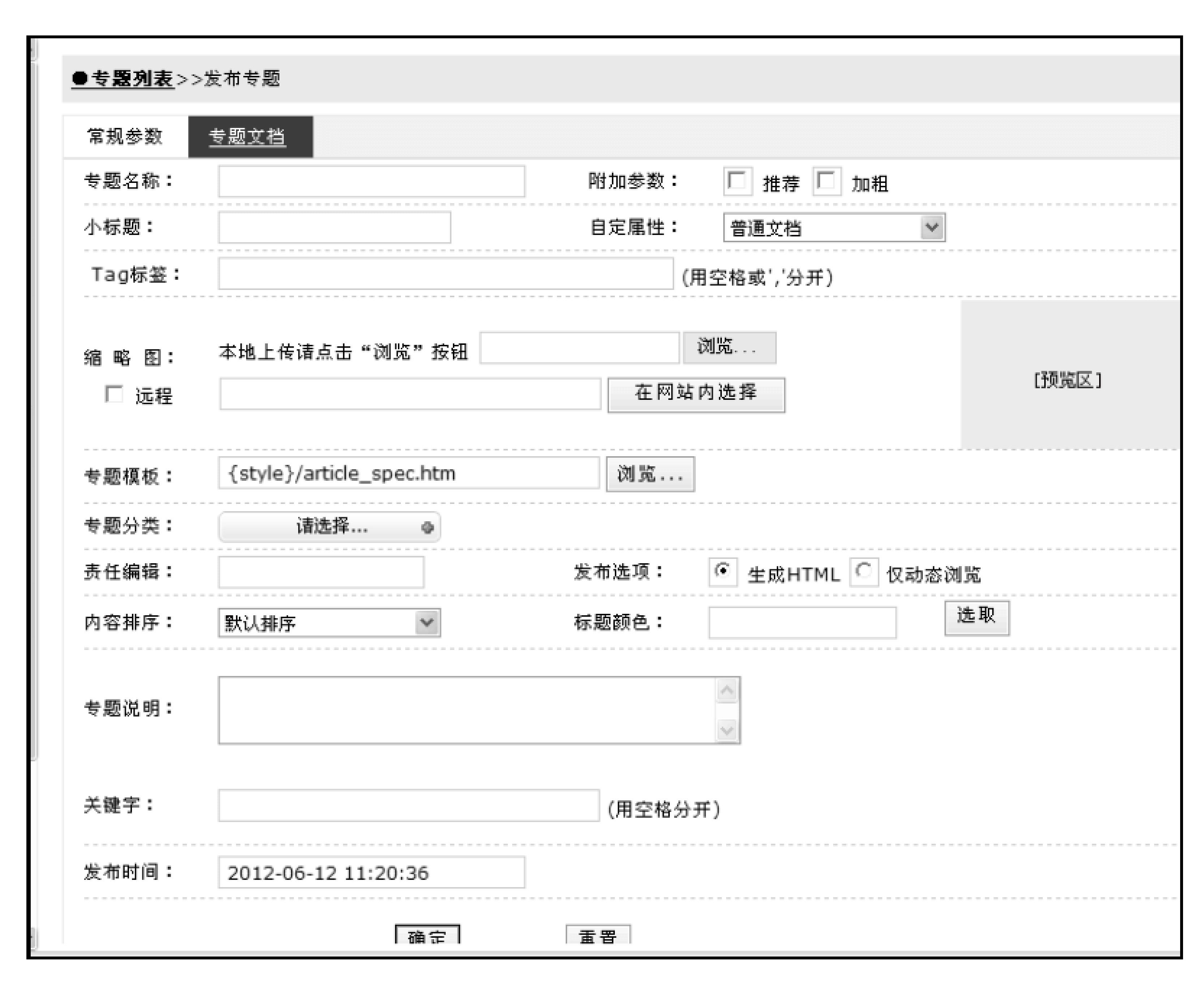


图 7-12　专题参数设定和文章设定

步骤 03　在专题的“常规参数”选项卡中，有专题名称、关键字、专题说明、专题 URL 结构设置等项目，下面进行简单介绍。

“专题名称”这个名称就是专题呈现出来的名字，在通常情况下，应该是三级关键词的名称，如图 7-13 所示。

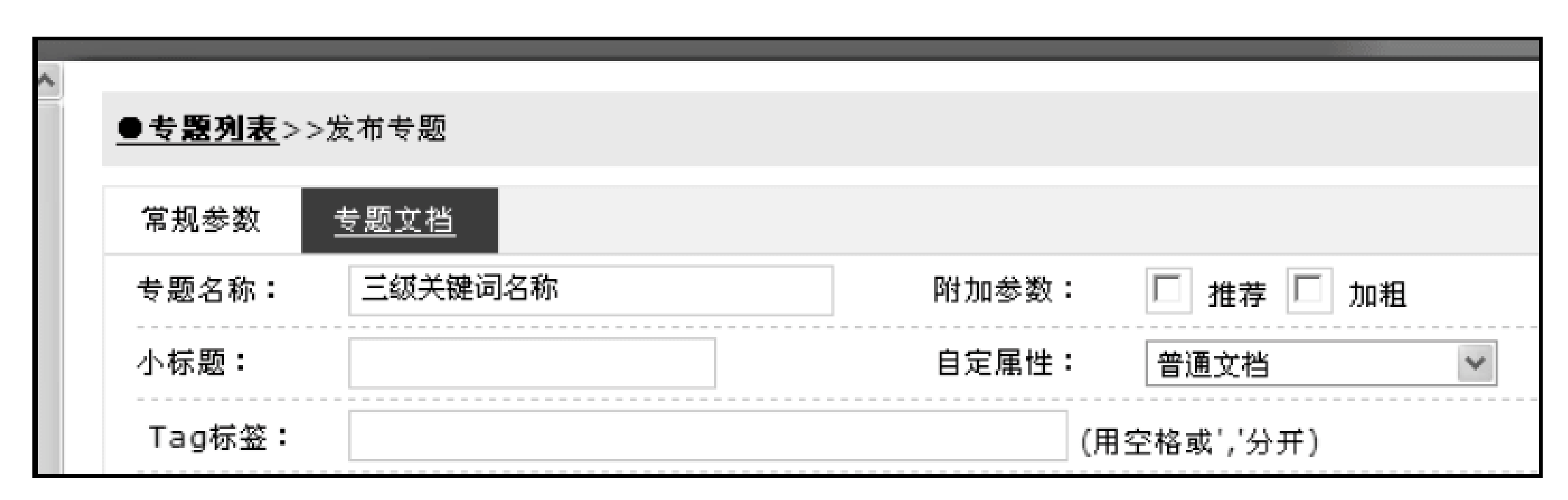


图 7-13 设置专题名称

“关键字”当前专题下的关键词内容，如果写不好，可以不写。不建议选择“自动获取”，如图 7-14 所示。



图 7-14 设置关键字

“专题说明”需要注意这个选项。很多人认为这个说明就是网页中的 description，实际上不是。专题说明是展示给用户直接浏览的，相当于栏目建设中的“栏目介绍”，所以专题说明一定要认真地规划内容，不然就失去了制作专题的意义，如图 7-15 所示。

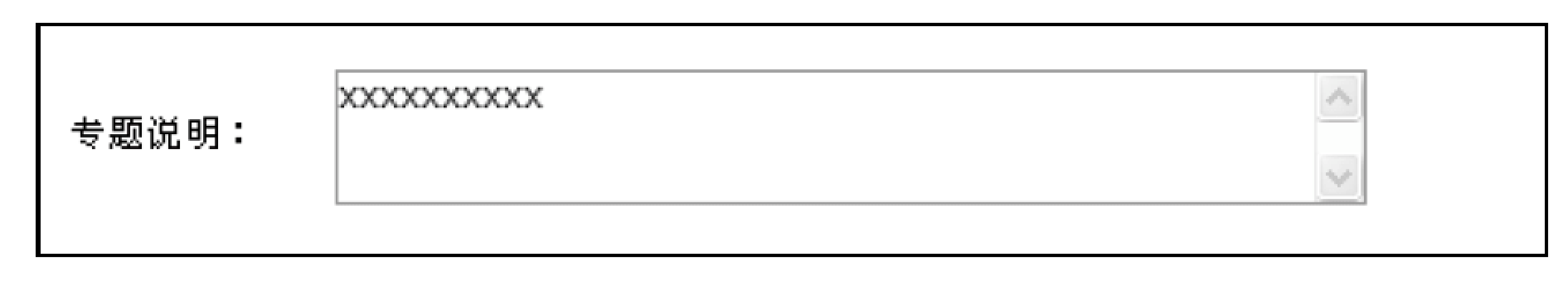


图 7-15 规划专题说明

“自定义文件名”这个选项主要用于规划专题物理结构和指定文件名，也就是确定专题存放的位置和专题所在页面的具体命名规范，一般比较实用的方法是放置在相关栏目下面，并以相对关键词的拼音进行命名，如图 7-16 所示。



图 7-16 自定义文件名

步骤 04 上面的介绍是关于专题的“常规参数”的，设定好之后，专题下面的文章也需要设置。在不同的 CMS 系统中，划定专题相关文章的方法略有不同，但是实质是一样的：都是选择和专题内容相关的文章来制作成列表，丰富当前专题，如图 7-17 所示。

●专题列表>>发布专题

常规参数　专题文档

专题节点列表：

1、文章列表用ID1,ID2,ID3这样形式分开，系统会自动排除不同节点的相同文章；

2、关于单条记录模板里的[field:fieldname /]标记的使用，请参考关于系统帮助关于 arclist 标记的说明；

3、节点ID是节点的唯一标识，允许在专题模板中用{dede:specnote id='标识'/}这样来表示单个节点。

4、正常的情况下，每个节点显示的是“节点文章列表”里的文档，如果你指定为“自动获取模式”，那么必须指定关键字和栏目ID。

节点 1 名称：　节点标识：

节点文章列表：　选择节点文章

文档来源：　按文章列表　自动获取文档　关键字：　(逗号分开) 栏目ID：0

节点布局：　列数：1　图片高：90　图片宽：120　标题长：60　简介长：160　文档数：40

单条记录的模板：

[field:title /]

节点容器模板：　system/channel/channel_spec_note.htm　浏览...

图 7-17　文章列表

通过上述的各个步骤，专题的建立就完成了。最好的专题建立作用是聚合不同栏目下的文章并抓取新的热点，需要优化者时刻关注业界动态。

三级关键词的结构规划一般是依靠小栏目和专题来进行的，比一、二级关键词规划要较为繁琐。下面进入数量最多、最繁琐，但是见效最快的长尾关键词网站结构规划。

7.7　做长尾关键词数据库

长尾关键词的最大特点就是优化难度低、数量大，从网站结构上说，长尾关键词就是网站结构金字塔的底端。体积足够大，层级却不高。通常情况下，长尾关键词的结构规划一般都会用内容页进行，不过需要注意的是，内容页除了承载长尾关键词外，还可以旁生出更多的长尾关键词。

7.7.1　长尾关键词库的建立

1. 什么是长尾关键词

通常情况下，网站上非主要的关键词，但是和网站主题有关联，具体说明某一问题的词或者短句，并且可以带来一定搜索流量的关键词，称为长尾关键词。长尾关键词的特点是比较长，一般由 2～3 个词、短语组成。长尾关键词往往存在于内容页面，除内容页的标题外，还可能存在于不容易被关注的网页内容中。

与主关键词相比，长尾关键词带来的搜索量比较少，并且很不稳定，这就决定了长尾关

键词无法像主关键词那样去重点优化。但是，长尾关键词带来的客户，转化为网站产品客户的概率比其他层级的关键词高很多，因为搜索长尾关键词的用户都有比较精确的需求。另外，对内容有一定质量的一个站点来说，长尾关键词可以为整个网站带来 20%～30%甚至更多的流量。

长尾关键词的构成很复杂，从搜索引擎优化者和网站站长的角度归结起来可以概括成以下 3 个方面。

- 与企业产品或者网站定位精确度高的词语。此类关键词针对的是有明确目标需求的搜索引擎引入用户，这类人群是对你所经营的网站产品、网站主题有着明确认知程度的人群，但是这部分人群只是网站流量的一小部分。
- 产品或者网站业务扩展出来的关键词。这是面向对你经营的网站或产品有着模糊概念的访问人群设计的关键词。
- 即将有可能成为搜索用户使用并且找到网站的关键词。这类关键词可能是业务周边产品的延伸词，或者是网站主题内容扩展出来的相关词。

举例来说，以“三亚旅游”这个关键词作为基本关键词进行长尾扩展，就可以大概知道什么是长尾关键词了。

针对性比较强的长尾关键词可能包括：“三亚蜈支洲”、“三亚蜈支洲旅游体会”等。

相关拓展的长尾关键词可能包括：“三亚宾馆织梦订”、“最低三亚旅游费用”等。

周边产品的延伸长尾关键词可能包括“北京到三亚机票预订”、“北京到三亚打折机票”、“三亚旅行社有哪些”、“三亚就餐推荐”等。

2. 如何获得长尾关键词

长尾关键词的获得是一个缓慢积累的过程，在通常情况下，长尾关键词不能像主关键词那样清晰、完善地进行罗列，需要通过日常网站的维护、流量统计报告等方式不断收集、不断扩充。

网站优化者在建设关键词库时，收集到的很多很长、内容上有一定相关性的词，可以当做长尾关键词使用。通过搜索引擎的相关搜索等功能，可以获得很多长尾关键词，如图 7-18 所示。

相关搜索	什么是长尾关键词	长尾关键词挖掘	长尾关键词挖掘工具	什么叫长尾关键词	长尾关键词查询
	长尾关键词工具	百度竞价长尾关键词	如何做长尾关键词	长尾关键词组合工具	长尾关键词举例

图 7-18　百度的相关搜索

通过各种关键词挖掘工具，可以获得很多长尾关键词，如图 7-19 所示。



序号	关键字	搜索量	收录数
1	狗狗	45300	100000000
2	狗狗书籍	30088	2160000
3	狗狗搜索	11109	3350000
4	狗狗下载	6900	3770000
5	狗狗影视	4256	2760000
6	大狗狗	3123	2620000
7	狗狗图片	2955	5970000
8	迅雷狗狗	1417	3090000
9	狗狗网	1217	8530000

图 7-19　关键词挖掘工具

在网站运行之后，通过流量统计报告，可以发现很多稀奇古怪的关键词来源，这些就是最直接的长尾关键词，如图 7-20 所示。

关键字	搜索次数	独立访客	IP	新独立访客	历史	查看引擎
白玉鸟	139	127	127	106	历史	点击查看
宠物狗大全	138	115	114	109	历史	点击查看
罗威纳论坛	136	43	43	33	历史	点击查看
卡斯罗	133	95	95	75	历史	点击查看
怎么分辨纯种泰...	131	121	121	99	历史	点击查看
比熊论坛	130	54	54	51	历史	点击查看
卡斯罗犬价格	123	89	88	70	历史	点击查看
中亚犬	120	97	97	73	历史	点击查看
银狐犬	116	104	103	95	历史	点击查看
狗狗图片大全	116	88	88	84	历史	点击查看
泰迪俱乐部	116	42	42	36	历史	点击查看
水晶虾图片	115	105	104	96	历史	点击查看
阿富汗猎犬	115	99	99	92	历史	点击查看
鹦鹉价格	114	110	109	102	历史	点击查看
斯坦福犬	112	89	89	74	历史	点击查看
猫	111	103	102	100	历史	点击查看
凯利蓝梗	110	83	82	69	历史	点击查看
狗狗图片	110	83	80	72	历史	点击查看
孔雀鱼论坛	110	75	75	66	历史	点击查看
八哥犬	109	96	96	80	历史	点击查看
宠物狗的种类	108	97	97	90	历史	点击查看
狗狗品种大全	108	89	89	86	历史	点击查看
狗狗品种图片	107	86	86	82	历史	点击查看

图 7-20　流量统计中的各种关键词

7.7.2　长尾关键词库的建立和维护

通过各种途径获得了长尾关键词以后，搜索引擎优化者应该建立一个长尾关键词库，具体建立方式因人而异。可以采用 Excel，也可以采用其他数据库——不过不建议用文档的方式建立长尾关键词库，因为长尾关键词即使对于大型网站来说，可能有几千甚至上万个，并且还在不断增加，普通文档格式不易维护。使用 Excel 建立的一个长尾关键词库，如图 7-21 所示。

	C	D	E	F	G	H	I	J
1	7天指数	相关页	首页数	广告	ER	EPJ	百度下拉框	
2	12917	200000	5	0	2	2.6	下拉	
3	9588	3330000	4	0	2	2.6	下拉	
4	7115	7290000	17	0	3	2	下拉	
5	6635	6880000	14	0	3	2	下拉	
6	6586	6300000	2	0	3	2.6	下拉	
7	6229	7750000	15	0	3	2	下拉	
8	6169	730000	4	0	3	2	下拉	
9	6112	12600000	11	0	3	2.6	下拉	
10	5820	2070000	4	0	2	2	下拉	
11	5907	347000	3	0	2	2		
12	5694	120000	3	0	2	2	下拉	
13	3972	13600000	19	4	3	1.5	下拉	
14	4511	1730000	15	1	3	2	下拉	
15	2467	1260000	6	0	3	2	下拉	
16	2084	32000	2	0	2	2		
17	2275	1280000	3	0	2	2	下拉	
18	1831	1620000	7	0	2	1.5	下拉	
19	1844	1530000	26	0	2	2.6	下拉	
20	2143	1480000	6	0	2	2.6	下拉	
21	1634	1820000	5	0	2	2	下拉	
22	1468	1350000	10	0	2	2	下拉	
23	1414	2180000	2	0	2	2		
24	1899	2300000	4	0	3	2	下拉	

图 7-21　使用 Excel 建立的关键词库

获得了长尾关键词库，又该如何对它们合理地进行布局规划呢？从网站结构规划的角度来说，最好的办法就是通过内容页。

7.7.3　内容页与长尾关键词的结构优化

所谓内容页，就是一个网站中最常出现的“文章页”。内容页的建设方式很简单，发布内容即可。但是内容该如何建设？长尾关键词应该怎么规划？这就非常繁琐了，也是检验一个搜索引擎优化者基本功的地方。

获得长尾关键词后，可以适当地将长尾关键词融入网站，以便获得更好的用户体验和更多的流量。最简单、常见的方法步骤如下。

步骤 01　针对每一个长尾关键词专门写一篇文章，发布一页内容。

步骤 02　在网站文章中涉及某长尾关键词时，加上链接，方便用户点击。

步骤 03　在和长尾关键词有关联的页面上，做“相关推荐”，将用户引入长尾关键词页面。

7.7.4　有效的长尾关键词布局规划

从优化理论来说，一个内容页应该包含 1～3 个长尾关键词，也就是说，站长在撰写内容时，应该围绕几个主要内容分支进行。不过在实际的网站建设中，很难严格地做到这一点，而可行的方式是自然、随意地融合长尾关键词。

举例来讲，如果优化者要针对一个名为“平和体质的表现”的长尾关键词做内容，不建议将这个内容页的标题就设定为“平和体质的表现”。应该考虑和这个词类似的、有密切关联或者有同一意义的其他相关词的优化，比如“平和体质的特征”等。结合两个长尾词得到最后的内容页标题可能就是“平和体质的表现和特征”、“平和体质的表现、特征”、“平和体质的表现以及特征”等。

采用这种方法规划长尾关键词，一方面是从逻辑、语意上需要这样做，这样做也更符合用户体验的需求；另一方面是因为网站内容并不一定有这么多的储备，也就是说，很难针对同义、近义的长尾关键词分别撰写高质量的文章。

总的来说，长尾关键词在网站建设、网站结构规划中有很重要的意义，需要搜索引擎优化者关注。但是长尾关键词比较难把握，需要细心和耐心才能让它发挥作用，养成做长尾关键词“厚积薄发”的习惯，对搜索引擎优化和流量吸引有莫大的好处。

网站布局非常重要，一定要特别注意如何更好的布局。

1. 网站标签的布局

所谓网站的是指整套布局好的 Title、Keywods、Ddescription，以便于简单清晰地告诉搜索引擎我们的网站是干什么的。

2. 网站分类布局

根据网站的性质来决定目录分类，可以根据价格、行业、产品类型、时间、区域、人群等方式进行操作。

3. 网站布局位置

网站布局一定要符合用户体验，很多网站的操作人员经常会为了突出网站的介绍从而浪费很多资源，如企业介绍、联系我们之类的，介绍和联系其实可以合二为一。另外如果相关简介的位置非常好，也很容易造成资源浪费，访客到我们网站是寻找自己需要的东西而不是来观看网站的企业信息，如何布局既能符合SEO又能兼顾用户体验，这里总结出了以下几点。

- 位置越靠前，用户越容易找到。
- 代码越靠前，搜索引擎越能快速抓取到。
- 越是靠前的信息越为重要，我们经常搜索任何关键字都会看到一个网页的标题和描述，就证明了这一点。

所以我们应该合理利用最为重要的网站位置，不要造成大量的资源浪费。如果觉得没有必要的信息或者不希望被搜索引擎抓取，那么就可以在网站的 Robots 里添加或者用 Nofollow 进行设置。

4. 网站布局内容

网站内容的布局也很重要，搜索引擎喜欢原创的，以及用户有需求但目前互联网上未曾出现的信息，一旦我们坚持这样的内容更新，那么会加快我们网站的收录速度，由此可见原创性的重要性。

5. 网站链接的布局

前面的事情做好之后，就开始进行链接的布局了，我们经常会说为什么百度收录突然减少？这是因为布局还不够好，内容的更新会造成老内容的信息过多或者老内容的质量不足被搜索引擎给删除。所以我们经常提到布局简单化，结构扁平化就是这个道理，因此需要我们做好网站的内部链接。

7.8 搜索引擎对重要页面的抓取

搜索引擎优化（SEO）是搜索引擎有好的收录网页的过程，适当的 SEO 有利于蜘蛛爬行你的网站，使你的内容符合搜索引擎算法，以便确认它与关键词高度相关。优化的目的就是使网页的内容便于蜘蛛理解，下面就以百度新收录的一篇内容来进行说明，笔者以一个完整的标题进行搜索，其目的是能反映出更多关键词在百度自然搜索的结果页面，能更好地说明搜索引擎是如何抓取你的网页的，如图 7-22 所示。



图 7-22　百度最常见的自然搜索结果列表

SEO 通常关注的是网页主体部分的自然搜索结果，百度和 Google 搜索结果页面大致相同，搜索结果页面通常为顶部、右侧、尾部均为广告链接，热门关键词会有广告，而笔者用这么长的词语就属于冷门词语，没有广告。

页面标题很重要，搜索引擎通常抓取的页面标题是 HTML 代码中的标题 title 部分，所以页面标题的写法对搜索引擎排名有重要的意义，通常页面标题 title 的优化格式为：文章标题→栏目标题→网站标题，它不一定只展示 description。

而第二行的描述部分通常会截取description 部分，但大多时候会抓取正文的前 200 字，但具体是显示 description 还是截取正文，这和用户搜索的关键词有关。

由于互联网中信息量极其庞大，为了向用户展示更多有价值的信息，搜索引擎会优先抓取每个网站中相对重要的页面（即权重较高的页面）。然而，搜索引擎是怎样发现这些重要页面的呢？根据重要页面链接指向的页面也可能是重要页面的思想，搜索引擎首先会从权重相对较高的页面（即源页面）出发跟踪其中的链接，从而抓取其他相对重要页面（即目标页面）中的链接，再抓取网站中其他相对重要的页面。

搜索引擎对页面中每个区域的重视程度是不一样的，同样的内容出现在页面中不同的区域，所起的作用也会存在很大的差异，这就涉及到页面重要区域的分布规律问题。

7.9 网站地图

网站地图，又称站点地图，它就是在一个页面中放置了网站上所有页面的链接。大多数访问者在网站上找不到自己所需要的信息时，可能会将网站地图作为一种补救措施。搜索引擎蜘蛛非常喜欢网站地图。

网站地图是一个网站所有链接的容器。网站地图可以方便搜索引擎蜘蛛抓取网站页面，通过抓取网站页面，清晰了解网站的构架。网站地图一般存放在 Robots 文件中，为搜索引擎蜘蛛指路，增加网站重要内容页面的收录。网站地图就是根据网站的结构、框架、内容生成的导航网页文件。网站地图可提高用户体验：它们为网站访问者指明方向，并帮助迷失的访问者找到他们想看的页面。

对于 SEO，网站地图有以下几个好处。

- 为搜索引擎蜘蛛提供可以浏览整个网站的链接。
- 为搜索引擎蜘蛛提供一些链接，指向动态页面或者采用其他方法比较难以到达的页面。
- 作为一种潜在的着陆页面，可以为搜索流量进行优化。
- 如果访问者试图访问网站所在域内并不存在的 URL，那么这个访问者就会被转到“无法找到文件”的错误页面，而网站地图可以作为该页面的“准”内容。

7.9.1 网站地图设计的技巧

网站地图它不仅要满足访问者的需求，还要取悦搜索引擎蜘蛛。因此在设计上需要注意以下几个方面。

1. 网站地图要包含最重要的页面

如果网站地图包含的链接过多，访问者浏览的时候就会迷失。因此如果网站页面总数超

过了 100 个的话，就需要挑选出最重要的页面。建议挑选以下页面放到网站的地图中去。

- 产品分类页面。
- 主要产品页面。
- FAQ 和帮助页面。
- 位于转化路径上的所有关键页面，访问者将从着陆页面出发，然后沿着这些页面实现转化。
- 访问量最大的前 10 个页面。
- 如果有站内搜索引擎的话，就挑选出从该搜索引擎出发点击次数最高的那些页面。

2. 网站地图布局要简洁，且要尽可能多的包含关键字

不要使用图片来做网站地图里的链接，这样蜘蛛就不能跟随了。一定要使用标准的 HTML 文本来做链接，链接中要包含尽可能多的目标关键字。如可以使用“无公害除草剂、杀虫剂和杀菌剂”来代替“我们的产品”作为标题。

3. 尽量在站点地图上增加文本说明

增加文本会给蜘蛛提供更加有索引价值的内容，以及有关内容的更多线索。

4. 在每个页面中都放置网站地图的链接

访问者一般会期望每个页面的底部都有一个指向网站地图的链接，你可以充分利用这一习惯。如果网站有搜索栏的话，则可以在这个搜索栏的附近增加一个指向网站地图的链接，甚至可以在搜索结果页面的某个固定位置放置网站地图的链接。

5. 确保网站地图里的每一个链接都是正确、有效的

如果在网站地图中出现的链接是坏链和死链，对搜索引擎的影响是非常不好的。如果链接比较少，你可以把所有的链接都点一遍，以确保每一个链接都是有效的。如果链接比较多，可以使用链接检查工具来检测。

6. 可以把 Sitemap 写进 Robots.txt 里

在搜索引擎蜘蛛进来抓取网页的时候，会首先查看 Robots.txt，如果首先把 Sitemap 写进 Robots.txt 里，那么在效率上会大大提高，从而获得 SEM 的好感。

7. 网页能够给访问者带来良好的体验尤为重要

网上有很多 Sitemap 地图生成的方法，比如在线生成、软件生成等，Sitemap 地图可以提交给各大搜索引擎，从而使搜索引擎更好地对网站页面进行收录，我们也可以通过Robots.txt 来告诉搜索引擎地图的位置。将制作好的网站地图上传至网站根目录下，最关键的是要把网站地图链接地址加入在 Robots 文件中以及做好网站地图在页面方便蜘蛛抓取的位置，一般把网站地图放在页眉和页脚位置。

7.9.2 普通 HTML 格式的网站地图

HTML 格式的网站地图根据网站结构特征制定，尽量把网站的功能结构和服务内容有条理地列出来。一般来说，网站首页有一个链接指向该格式的网站地图。

XML Sitemap 通常称为 Sitemap（首字母大写 S）。简单来讲，Sitemap 就是网站上链接的列表。制作 Sitemap 并提交给搜索引擎可以使网站的内容完全被收录，包括隐藏得比较深的页面，这是一个网站与搜索引擎对话的好方式。

因为每个搜索引擎主要识别地图格式效果不同，建议分别采用以下格式。

- 百度：建议使用 HTML 格式的网站地图。
- Google：建议使用 XML 格式的网站地图。
- Yahoo！：建议使用 TXT 格式的网站地图。

7.9.3 网站地图的注意事项

建议经常更新网站地图，便于培养搜索引擎蜘蛛爬行的频率。经常有新的地图内容生成，长期以来，蜘蛛就会更关注，培养蜘蛛的爬行规则。这样网站内容能更快地被搜索引擎抓取收录，网站内容也可以早日被搜索引擎检索。

1. 全部链接真实有效

地图的主要目的是方便搜索引擎蜘蛛抓取的，如果地图存在死链或坏链，会影响网站在搜索引擎中网站权重的，所以要仔细检查有无错误的链接地址，在提交前通过站长工具，检查网站的链接是否有效。

2. 简化网站地图

网站地图不要出现重复的链接，要采用标准 W3C 格式的地图文件，布局要简洁、清晰，如果地图是内容式地图，每页不要超过 100 个内容链接，采用分页的形式，逐一开来，这样方便搜索引擎蜘蛛逐页爬行。

3. 网站地图的多样性

网站地图不仅仅是给搜索引擎看的，而 SEO 的根本目的也是方便访问者，所以网站地图在兼顾搜索引擎的同时也要兼顾访问者。我们通常为一个网站建 3 个网站地图。Sitemap.htm 页面精美，简洁大方，让访问者便于找到目标页面的同时也能愉悦心情。认真研究自己的网站，把重要的页面标注出来，在不需要纳入的页面加上 No Follow 这样更有利于搜索引擎辨别。Urllist.txt 或者 Robots.txt 如果方便最好做一下，Yahoo！等搜索引擎比较认可，谷歌也有这个项目。

7.10 小结

很多站点在开始策划网站的时候都没有把网站布局思考透彻，然后就经常对自己的网站改版，也导致了网站的一些负面作用，为搜索引擎蜘蛛所不喜。因为修改太多，导致搜索引擎不给最新快照、收录较慢，这几点也是笔者亲身体验的一些总结，希望对各位站长和 SEO 有所启示和帮助。

第 8 章　链接策略

链接（即超链接）指从一个网页指向另外一个目标的连接关系，这个目标可以是一个网页，也可以是同一个网页的不同位置，还可以是一个图片、电子邮件地址或文件，甚至可以是一个应用程序。在网站优化工作中，链接的优化也很重要。本章将介绍链接的策略知识，包括站内链接，内部链接，外部链接及增加外部链接等。

8.1 站内链接

现在的用户进入网站有时候会浏览同一个网站下面不同类的内容，这个时候他往往就会想到导航或者分类等，这些东西也是站内链接，现在搜索引擎对于用户体验是非常关注的，所以站内链接非常重要。站内链接优化有以下好处。

（1）可以提高页面蜘蛛抓取效率

我们都知道蜘蛛抓取网页分为广度优先和深度优先，不管是哪种方式靠的都是站内链接的合理格局，所以只有提高了蜘蛛抓取页面的频率，才能保证网站能够具有更好的收录。

（2）能够增加网站的收录数量

其实这只是对上面一条的补充，提高了页面蜘蛛抓取率后，只要内容还行，质量不错，那么它就会收录你的网站，因此网站收录数量自然而然的提高了。

（3）使得网站的权重得到提升

其实搜索引擎对每个网站页面给的权重都是一样高，特别是谷歌，它给予每个页面的权重是一样的，而每个页面都能赋予其他页面权重。我们每天辛苦的去做外链，就是在不同的网页页面上留下我们的链接，让这个网页分点权重给我们。同样的，我们网站内部网页也是能给同一个网站下面的不同网页带来权重，所以这个站内链接又可以传递权重，提高关键词排名，给整个网站的权重带来一个整体的提高。

8.1.1 链接决策

链接是决定页面权重最主要的因素，分内部链接及外部链接。本章将从内、外部链接的角度出发，讲述链接对页面权重及相关性的影响，以及介绍如何合理、有效地增加外部链接。

1. 源页面与目标页面

在链接关系中，我们称存放链接的页面为源页面，被链接的页面为目标页面。如页面 A 中存在指向页面 B 的链接，则页面 A 为源页面，而页面 B 为目标页面。

2. 链接与投票机制

在搜索引擎里，链接反映的是页面间的信任关系。也就是说，如果某个页面中存在链接指向另外一个页面，则表示该页面对被链接的页面是信任的。

在搜索引擎优化里，利用“投票机制”来表达页面间的这种信任关系。如在页面 http://www.123pet.com.cn/action-news.html 中存在链接指向页面。http://www.123pet.com.cn/ask/，则说明页面/action-news.html 对/ask/是信任的，从而投了它一票。

3. 链接对象

在网页中，超链接的对象可以是文本、图片或者多媒体文件。当访问者点击链接对象后，链接目标将按照指定的方式打开或运行。

4. 链接分类

按照链接使用的对象可分为“文本链接”、“图片链接”及“多媒体文件链接”；按照链接的指向可分为“导入链接”与“导出链接”；而按照链接的范围分为“内部链接”和“外部链接”。

（1）文本链接

文本链接指使用文本作为链接对象的链接。如：宠物网就是以文本“宠物网”作为链接对象指向目标地址：http://www. 123pet.com.cn 的。

在对象是文本的链接里，我们称文本内容为锚文本。如某页面中存在这样的链接：宠物网，则“宠物网”就是该链接中的锚文本。

（2）图片链接

图片链接指使用图片作为链接对象的链接。如： <img src="http://www. 123pet.com.cn/images/logo.png"> 就是以图片 logo.gif 作为链接对象指向目标地址：http://www. 123pet.com.cn 的。

（3）多媒体文件链接

多媒体文件链接指使用多媒体文件作为链接对象的链接。在多媒体文件链接里，链接信息包含在多媒体文件中，当用户点击该多媒体文件时就会打开相应的目标内容。

如以下代码，表示把链接信息：http://www. 123pet.com.cn 放在 Flash 文件里。当用户点击该 Flash 文件时，就会打开目标地址：http://www. 123pet.com.cn。

```
on （release） {
getURL（"http://www. 123pet.com.cn", "_blank"）;
}
```

（4）导入与导出链接

假设 A 和 B 是两个不同的页面，页面 A 中存在链接指向页面 B（即 A→B），则我们称该链接是页面 B 的导入链接（也就是我们常说的反向链接），页面 A 的导出链接，即页面 B 得到了页面 A 的投票，页面 A 给页面 B 投了一票。

（5）内部链接

内部链接指网站内部页面间的链接关系。

（6）外部链接

外部链接是指与外部网站页面间的链接关系。

8.1.2　站内链接的意义

站在搜索引擎的角度，链接会直接影响目标页面相关性及权重。链接对目标页面相关性的影响主要取决于链接的对象及内容。前面介绍的三种链接对象，在提高页面相关性方面影响最大的是文本，其次是图片，最小是多媒体文件。接下来，将从多个角度剖析链接对目标页面相关性及权重的影响。

（1）文本链接

文本链接在提高目标页面相关性方面所起的作用最大。因为，文本链接可以通过锚文本直接、有效地表达目标页面的主题。

如在某页面中存在这样的一个链接：宠物网，该链接通过锚文本“宠物网”表达了目标页面的主题。

在使用文本作为链接对象时，应该尽量使用与目标页面主题相关的关键字作为锚文本。

①在同一页面上分别使用以下两个不同锚文本的链接。

链接 1：pett

链接 2：宠物

在提高网站“www. 123pet.com.cn”与关键字“宠物”相关性方面，链接 2 所起的作用远远大于链接 1。

②在同一页面中，即使链接的锚文本相同，目标页面与锚文本间的相关性也会由于锚文本样式的不同而有所差异，如下所示。

链接 1：宠物网

链接 2：宠物网

在这两种链接方式中，链接 2 在提高目标页面与锚文本“宠物网”相关性方面所起的作用要比链接 1 大。

（2）图片链接

由于搜索引擎并不能识别图片中的文本内容，图片链接在提高页面相关性方面所起的作用几乎是可以忽略的。但是，在使用图片作为链接对象时，可以通过设置图片的<alt>标签属性值来表达目标页面的主题。例如：

- <img
- src="www. 123pet.com.cn/images/logo.png.
- gif" alt="宠物">

在上例中，通过设置<alt>标签属性值来提高网站“www. 123pet.com.cn”与关键字“宠物”的相关性。但是，这种间接的表达方式所起的作用远不如文本链接。

（3）多媒体文件链接

搜索引擎解析多媒体文件的几率是非常低的，某些搜索引擎甚至会忽略多媒体文件（如百度及雅虎）。因此，多媒体文件链接在提高目标页面相关性方面所起的作用是可以忽略的。

综上所述，对于重要的目标页面，应该优先使用具有特别样式的文本作为链接对象，且锚文本应该采用与目标页面主题相关的关键字，这样就可以在最大的程度上提高目标页面的相关性。

8.1.3 站内链接的优化

其实站内链接优化这样的文章在网上很多，而且大体都相同，因为做网站优化 SEO 主要是去迎合搜索引擎的习惯来进行优化的，所以对于站内链接的可控因素，大体上做得都差不多，这里总结的经验如下。

（1）给网站建立地图

这里主要介绍思路，其实地图的作用主要是提高搜索引擎的工作效率，让网站的每个页面都呈现在搜索引擎面前，这样抓取的几率也大。对于网站地图的格式有 HTML、XML、TXT 等，建议做 HTML 这样更符合百度的习惯。

（2）导航

做导航时候要注意尽量不用 Flash、图片、js 等，因为目前的搜索引擎是抓取不到这些的，不过如果非要用图片做的话可以加入<alt>标签，导航其实可以分为很多种，主导航、副导航（除了主导航外的其他导航统一叫副导航）、路径导航等，如图 8-1、图 8-2 所示。



图 8-1　主导航



图 8-2　路径导航

（3）站内锚文本

即包含超链接的文字，锚文本可以传递权重、提高用户体验。

（4）相关文章

如上一篇、下一篇、相关文章或者最近发表等，这里其实也是增加了用户体验和曝光率，提高抓取率。

（5）点击次数

就是网站最底层的文章从首页点进来的点击次数，最好不要超过 3 次，一般的 3 次就是非常好的了，对于大型网站肯定超过 3 次，这里主要针对一般的网站，所以需要在网站正式推出之前就在结构上面设计好，以免后期进行改动。

（6）权重合理分配

一般来说，我们是给首页赋予最高的权重，栏目次之，最后是内容页。但发现很多网站只是给了首页外链或者内链，而内容页或者栏目页根本不管它，这是一个很大的误区思想，其实都要兼顾，这样网站才能良性发展。

（7）其他

站内链接需要注意的东西还有很多，给大家一个思路，多去一些大型网站、做了整体SEO 的知名网站、排名好的网站，去观察这些网站的站内链接是如何做的，善于观察和发现才是我们前进的好兆头。

8.2 内部链接

内部链接指网站内部页面间的链接关系，反映了网站内部页面间的信任关系。内部链接除了直接决定网站的逻辑结构，影响搜索引擎对网站页面的收录外，还影响网站中每个页面的权重及相关性。

8.2.1 内部投票机制

在网站内部，如果某一页面中存在链接指向另外一个页面，即表示该页面对于被链接页面是信任的，从而投了它一票。通过分析内部页面间的链接关系，搜索引擎可以从中筛选出相对重要的页面。

例如在同一个网站里，页面 A 存在链接指向页面 B，即表示页面 A 对页面 B 是信任的，因而给页面 B 投了一票。反过来，页面 B 得到了页面 A 的投票。

搜索引擎可以根据网站内部页面间的链接关系统计出页面的得票数，从而计算出每个页面的内部链接权重。在网站内部投票中，页面得到的投票越多，其重要性就越大，内部链接权重也就越高。因此，在对网站进行优化时，应该让网站中相对重要的页面得到更多的内部导入链接。

由于在提高目标页面与关键字相关性方面，文本链接所起的作用最大，因此，在规划网站内部链接时，应该优先使用文本作为链接对象，且锚文本应该采用与目标页面主题相关的关键字。

8.2.2 内部链接数量

搜索引擎对每个页面的内部链接数量是有一定限制的，如果页面中的内部链接的数量超过限制，则搜索引擎就可能会忽略该页面，或者忽略该页面中超出限制的那部分链接所指向的目标页面。因此，在规划页面时要尽量把页面的内部链接数量控制在合理的范围。

如 Google 就明确要求页面上的内部链接数要限制在 100 个以内。如果某页面的内部链接数量超过 100，Google 就有可能会忽略该页面，或者忽略 100 个以后的那部分链接指向的目标页面。Google 会有以下两种处理方式，如图 8-3 所示。

链接 1
链接 2
链接 3
……
链接 100
链接 101
……
链接 N

图 8-3　页面链接示意图

（1）忽略该页面。

（2）忽略该页面中“链接 101”及其后的链接指向的目标页面。

8.2.3　内部链接意义

由于网站所有者不但可以控制网站内部链接的对象，还能随意调整每个页面的导入、导出链接数量，从而控制内部页面的权重与相关性。相对于外部链接，搜索引擎对内部链接的重视程度较低，即内部链接在决定页面权重及相关性方面所起的作用较低（请参考 8.4 节）。

8.3 外部链接

外部链接是指本网站以外的链接，表达的是网站间的链接关系，反映了网站间的信任关系。与内部链接相反，外部链接具有不可操控性，即网站所有者不能通过正规的手段操控本站以外的页面的导入（导出）链接数、链接对象及链接目标。如你是网站 A 的所有者，你是控制不了网站 B 的页面上的导入（导出）链接数、链接对象及链接目标的。因此，网站所有者并不能操控页面的外部链接权重。

8.3.1　外部投票机制

外部投票机制表达的是网站间页面的信任关系。如网站 A 的页面 A-1 上存在链接指向网站 B 的页面 B-1，则表明网站 A 对于网站 B 是信任的，从而投了它一票。

同样，搜索引擎可以根据网站间页面的链接关系统计出每个网站的每个页面得到的外部投票数，从而计算出页面的外部链接权重。在外部投票机制中，页面得到的投票越多，其重要性就越大，外部链接权重也就越高。而且，由于外部链接存在不可操控性，通过外部投票

机制“推选”出来的页面，其重要性相对更高。

在提高目标页面相关性方面，也应该优先选择文本作为链接对象，且锚文本应使用与目标页面主题相关的关键字。此外，在选择外部链接时，还要考虑外部链接所在源页面的主题，即网站相关性的问题。

外部链接关系中，如果链接双方页面的主题是相同或相近的，则该链接在提高目标页面相关性方面所起的作用是非常大的。

8.3.2　外部链接数量

同样，如果忽略链接在源页面中的位置及存在时间，所有外部链接指向的目标页面也是以平均方式继承源页面的权重。因此，如果页面中存在过多的外部链接，不但会减少外部目标页面继承的权重，也会给网站内部页面带来致命的打击，甚至被搜索引擎视为垃圾链接页面。

例如很多网站的友情链接页面，动辄就是几百个外部链接。对于这样的页面，搜索引擎会把它当作是垃圾链接页面。

一般情况下，页面的外部链接数在 100 个以内是合法的。但综合考虑内、外部目标页面的权重继承等问题，在友情链接页面中，外部链接数量要控制在 40 个以内；而其他重要页面的外部链接数量应该控制在 20 个以内，如网站首页等。

8.3.3　外部链接的作用

（1）提升网站 PR 值

友情链接的质量好对提升 PR 值有很大的帮助，PR 值是传递的，这个就是为什么如果你的 PR 比对方低，人家都不愿跟你交换的原因，人都是自私的，都希望得到好处，虽然多次提到 PR 对百度的作用不大，但是在交换友情中很多站长还是会把这个因素考虑进去的。很多时候我们就因为这个，错失了优质链接。

（2）提高网站关键字的排名

一般做友情链接的时候，都会给别人提供一个锚文本，往往这个锚文本就是我们站点的一个关键词，友情链接是最好的外链之一，这个与我们去论坛，在博客做签名的作用是一样的，目的就是为了提高这个关键词的排名。

（3）带动网站的快照

蜘蛛是通过链接来抓取的，一个快照新的站点说明是受到蜘蛛喜爱的，这样我们和这些快照新的站点做友情链接，这样等蜘蛛抓取他网站的时候，也会通过连接来到我们的站点，就可以加快更新网站的快照。

（4）提高网站权重

来自首页的链接越多证明网站的权重越好（相关性越高，说明链接越好），而友情链接一般都是来自首页的链接，要是别人给我们做全站链接，那更好，给的权重就越好。所以友情链接站点导出链接越少，越对我们有利，因为人少分的东西也就越多。

（5）提高网站的访问量，减少网站跳出率。

我们和相关性的站点做友情链接，这样还是有流量进来的，只是通过这个进来的流量很少，但也算是额外来的流量。网站跳出率越低，证明我们站点越好，搜索引擎是从用户体验出来，所以说友情链接的作用还是很大的。

8.3.4　外部链接的意义

由于外部链接存在不可操控性，外部链接不管是在提高页面权重还是相关性方面所起的作用都远大于内部链接。

通过内部链接，搜索引擎可以发现网站中的新页面；而通过外部链接，搜索引擎几乎可以发现整个互联网。

8.4　增加外部链接

链接，特别是外部链接，在决定页面权重及相关性方面起着至关重要的作用。因此，为了提高页面的权重及相关性，我们应该让网站中相对重要的页面得到更多高质量的外部链接，如网站的首页。

寻找高质量外部链接是一项十分艰巨的任务，耗时长（搜索引擎优化 80%以上的时间都是耗费在寻找高质量外部链接上）、工作过程乏味。下面将介绍几种有效增加高质量外部链接的方法，分别是“分类目录”、“交换链接”及使用“链接诱饵”。

外链知识每个做网站的朋友都很了解，这里就不多讲了。那就进入正题，来看看给网站增加外链的方法和增加外链需注意些什么？

8.4.1　增加外链的途径和注意事项

1. 增加外链的途径

（1）可以找朋友的网站并加上你的链接。如果朋友愿意，单向链接是最好的。

（2）在 Google 和百度查询你网站的主题。如你是做游戏网站的，那就在 Google 和百度

中查询游戏，看看排名第一页的网站的反向链接，然后联系这些排名靠前的链接的网站跟你交换链接，或出钱跟他们做单向链接。

（3）查询排名好的大站，定期到他们站上发表一些原创文章。文章中巧妙地带上你网站的地址。

（4）去一些大型 Blog，尤其是已经被搜索引擎收录的个人 Blog，给他们评论时加上你的网站链接，同时也注意查找与你网站主题内容相关的 Blog 评论。注意，不相关的外链只会变成劣质外链。

（5）如果你有不错的技术，且你网站全部是原创的话，把你的网站内容输出成 RSS，提交到 Google 和百度的新闻中心。这样，就会有上万 IP 来你网站，也会被 N 多网站引用，因为好多网站都是引用 Google 和百度的 RSS 源，包括 IGoogle（Google 个性化主页）。想想 IGoogle 一天的使用量，可想而知给你带来的流量。当然要记住一定要保证你网站输出的 RSS 是原创内容，不然 Google 和百度不一定会青睐。

（6）将网站提交给开放目录 DMOZ（dmoz.org），DMOZ 收录网站的时间较长，可能需要 1 个月甚至 1 年时间，但收录后效果非常明显。

（7）在百度知道等问答系统，搜索你网站的关键字，回答相关的问题并留下你网站网址，注意一天不要回答太多，一般要超过 5 个，不然你的 ID 就会被封。

（8）如果你经常上论坛的话，可以把论坛的个性签名改成你网站的关键字链接。

2. 增加外链的注意事项

（1）增加外链在短时间不是越多越好，每天一条就够了，不然只会让搜索引擎注意到你，最后会封杀了你的网站，所以坚持才是重点。

（2）跟别人交换友情链接时一定要用 site：对方网站网址，检查对方网站在 Google 和百度是否收录，如果未收录，请千万别加他的友情链接，可能这个站是被搜索引擎封杀过的，你加上他的链接的话就会连累你的站。

（3）交换友情链接时链接锚文字一定要用你网站的关键字，如“金山打字通”中的金山打字通就是链接锚文字，这样作用才大，你在帮对方加友情链接时一定要用对方相关的链接锚文字。如对方是做游戏站的，却让你链接锚文字用化妆品，这样你绝对不能给他做，不然搜索引擎认为你作弊。

（4）一个页面上最多放 20 个友情链接，如果有更多的，做个更多友情链接页面做链接，包括对方给你做友情链接时，请不要做在链接超过 100 个的页面上，这样对你 PR 值的提升没有作用。

（5）如果对方给你的友情链接页面放在 IFRAME 框架里，不要跟他交换友情链接，因为搜索引擎根本抓不到 IFRAME 里的内容。

（6）如果对方网站给你做的友情链接目录下有 Robots.txt，且有 Disallow，还有友情链接的目录名，说明对方告诉蜘蛛不要去索引放在有友情链接目录下的任何文件。不要跟对方交换友情链接，检查有没有 Robots.txt，（在对方网址后加“/Robots.txt”就可看到），如果没有，就说明他没定义 Robots.txt，那就没关系了。

（7）有很多站点都在做“链接页”，如 links.htm、links.asp、links.cgo 等，如果搜索引擎看到这个，不仅可能把它记做一个内页链接，更有可能忽略不记，所以不要在这种页上交换友情链接。

（8）一定要定期检测你网站的友情链接。

（9）外链的位置也是一个很重要的因素。你的链接最好放在对方网站的头区（主导航区），主区（内容正文内容里），尾区（友情链接、次导航 版权及备案声明）。其中主区正文内容效果最佳，其次主区，最后到尾区，大部分搜索引擎已经降低了友情链接对排名的影响，这里的反向链接也只能算是劣质链接。一个主区正文链接可以抵上百个尾区劣质链接，所以前面讲到要在内容关键字上加链接就是这个原因。

（10）做友情链接时一定要用文字链接，不要用图片链接。因为搜索引擎抓对方网站时根本抓不到图片链接，对你一点用处都没有。

8.4.2 增加高质量链接的方法

一个好的网站，除了网站内部结构合理外，外链也是一个非常重要的因素，因为外链是提高网站 PR 值很重要的一个原因，而 PR 值又直接影响到网站的排名。

外链就是其他网站链向自己站点的链接。查询一个网站有多少外链，可在 Google 搜索框输“link：你网站网址”，如查询笔者的网站：link:www.123pet.com.cn 就可查询到外部网站链接到本站的外链数了。下面来介绍几种常用的增加外链方法。

1. 分类目录

分类目录是指通过人工的方式收集网站，并把这些具有一定价值的网站资源按照主题进行整理和组织，之后存放到相应的目录下，从而形成网站的分类目录体系。由于分类目录是由人工编辑而成，因此，又称为人工分类目录。

既然分类目录是由人工组建，就难免会受到主观因素的影响。然而，也正是因为受到主观因素的影响，分类目录里网站的质量才能得到保证。搜索引擎非常重视高质量的分类目录。

但对于低质量的分类目录，搜索引擎不但不重视，甚至会把它纳入垃圾链接制造厂。一旦被判为垃圾链接制造厂，不管是目录本身还是与该目录存在链接关系的网站都会受到相应的惩罚，如常见的自助链接系统。

目前，几大主流搜索引擎都有其特别重视的分类目录，即网站如果登录了搜索引擎特别重视的分类目录，则该站在对应的搜索引擎中就可以得到更高的权重。如 Google 比较重视 DMOZ，雅虎更关注自身的分类目录，而百度则更倾向于自家的 Hao123。

分类目录的链接具有单向性、高质量等特点，即成功登录分类目录的网站只是单方面继承分类目录中相应页面的权重，而不需要与分类目录分享网站自身的权重。

（1）简体中文目录

简体中文环境里从来就不缺分类目录，中国数以十万计的网址站就是最常见的分类目录。在简体中文环境里，常见的高质量分类目录如下所示。

- hao123：www.hao123.com
- 265：www.265.com
- 搜狗：www.sogou.com/dir/
- 中国雅虎：site.yahoo.com.cn
- 百度：site.baidu.com
- 谷歌：daohang.google.cn

其中，最重要的要数“hao123”及“265”。因为，目前国内很多网址导航站都是复制“hao123”及“265”这两个网站，登录了这两个网站就相当于登录了数以万计的网址站。

（2）英文目录

英文环境中，分类目录的数量也不少，只是高质量的分类目录一般需要付费才能登录，而普通的分类目录则以链接交换作为登录条件。对于那些无条件即可登录的分类目录，个人认为弊多利少。下面是两个较为著名的人工分类目录。

- ODP：www.dmoz.org
- 雅虎：directory.yahoo.com

其中，Google 对于登录 ODP 的网站会刮目相看，而雅虎则更加重视那些已经登录自家分类目录的网站。

（3）寻找分类目录

上面列出的只是几个中英文里较为著名的分类目录，俗话说：“授人以鱼不如授人以渔”。下面将介绍几个寻找分类目录的方法，希望能帮助读者寻找到更多免费的高质量分类目录。

①搜索

我们可以在 Google、雅虎及百度的相关关键字搜索结果中寻找分类目录。如可以以“分类目录”或“网址导航”作为关键字进行搜索，再在搜索结果中寻找分类目录。为了进一步缩小寻找范围，还可以结合与网站主题相关的关键字进行搜索，如“宠物分类目录”、“宠物网址导航”等。

此外，还可以在相关搜索中寻找分类目录，即点击相关搜索中的关键字寻找相应的分类目录，如图 8-4 所示。



图 8-4　宠物网相关导航

在英文语言环境中，可以使用同样的方式寻找分类目录。只是搜索的关键字会有所不同，如“directory”或“keyword+directory”等。

在中文语言环境中，除了在包含“分类目录”的关键字搜索结果中寻找分类目录外，还可以在包含“友情链接”或“交换链接”的关键字搜索结果中寻找。而在英文语言环境里，常在包含“directory”或“link exchange”的关键字搜索结果中寻找。

②在分类目录中再次寻找

除了利用搜索引擎寻找分类目录外，还可以在已寻找到的分类目录中再次寻找。如可以在Yahoo!分类目录里的“Search Engines and Directories”目录中寻找其他高质量的分类目录。

2. 交换链接

交换链接是指链接双方都在各自的网站上添加对方网站链接信息的行为。与分类目录不同，交换链接是建立在公平、互利的基础上（即交换链接双方的网站在重要性方面是相当的）。因此，在交换链接时，需要注意以下几点。

（1）网站相关性

在交换链接时，应该优先考虑与本站主题相同或相近的网站。如某网站的主题是“Flash小游戏”，则在交换链接时，应该优先考虑主题与“Flash”或者“小游戏”相关的网站，而非是毫不相关的网站。

（2）网站质量

以 Google PR 值衡量一个网站的质量，既简单又形象。但是，如果新建站点的 PR 值尚未更新，就得不到该站点的 PR 值信息。此时，可以利用网站的页面包含数作为衡量网站质量的依据。如某网站 PR 值是 0，但是页面包含数是 123 或 456，这样的网站的质量也是非常高的。

（3）导出链接数量

交换链接的页面上外部导出链接数越多（即链接伙伴越多），不管是网站内部页面还是其他链接伙伴能继承到的权重就会越少。因此，在选择链接伙伴时，应该优先选择外部导出链接数较少的网站。如外部导出链接在 20 个以内的页面。

（4）寻找友情链接

要交换友情链接，首先要有交换链接的对象，即链接伙伴。通过搜索引擎，可以在短时

间内找到大量的潜在链接伙伴。通常，可以以网站主题的名称作为关键字在搜索引擎中进行搜索，再在搜索结果中寻找潜在链接伙伴。如某网站的主题是“宠物”，则可以在“宠物”的搜索结果中寻找潜在链接伙伴。

我们还可以利用搜索引擎的相关搜索功能寻找潜在链接伙伴。如“宠物”的相关关键字包括“宠物网”、“宠物狗”、“宠物论坛”，则可以在这些相关关键字的搜索结果中寻找潜在链接伙伴，如图 8-5 所示。

相关搜索	qq宠物论坛	宠物鸟论坛	cpn宠物论坛	中国宠物鸟论坛	广州宠物论坛
	长沙宠物论坛	上海宠物论坛	柳州宠物论坛	中国宠物蚂蚁论坛	天涯宠物论坛

图 8-5　宠物论坛相关搜索

此外，还可以在“关键字+空格+交换链接/友情链接”的搜索结果中寻找潜在链接伙伴。其中，关键字可以是网站的主题名称，也可以是与网站主题名称相关的词语。如“小游戏友情链接”或者“迷你小游戏友情链接”等。

（5）分类目录

除了搜索引擎外，还可以通过分类目录寻找高质量的潜在链接伙伴。如可以在主题相同或相近的目录里寻找潜在链接伙伴，如图 8-6 所示。



图 8-6　宠物类网站目录

在分类目录里，可以在更短的时间内寻找到更多高质量的潜在链接伙伴。因为，分类目录中的每一个网站都是经过人工审核的，所以质量相对较高。但分类目录中网站的数量非常有限，且大部分是著名网站，交换链接的机会并不大。

（6）链接伙伴

除了通过搜索引擎及分类目录寻找潜在链接伙伴外，还可以从本站（或其他相关网站）链接伙伴的链接伙伴中寻找。如 A 是你的网站，B 是你的链接伙伴，C 是 B 的链接伙伴，则

你可以通过 B 找到 C，再通过 C 找到其他潜在链接伙伴。

可以在自己的链接伙伴 B 中找到潜在链接伙伴 C 至 C-n，再在 C 中找到其他潜在链接伙伴 N 至 N-n。这样通过链接伙伴一层一层地寻找，就可以得到数量庞大的潜在链接伙伴，如图 8-7 所示。

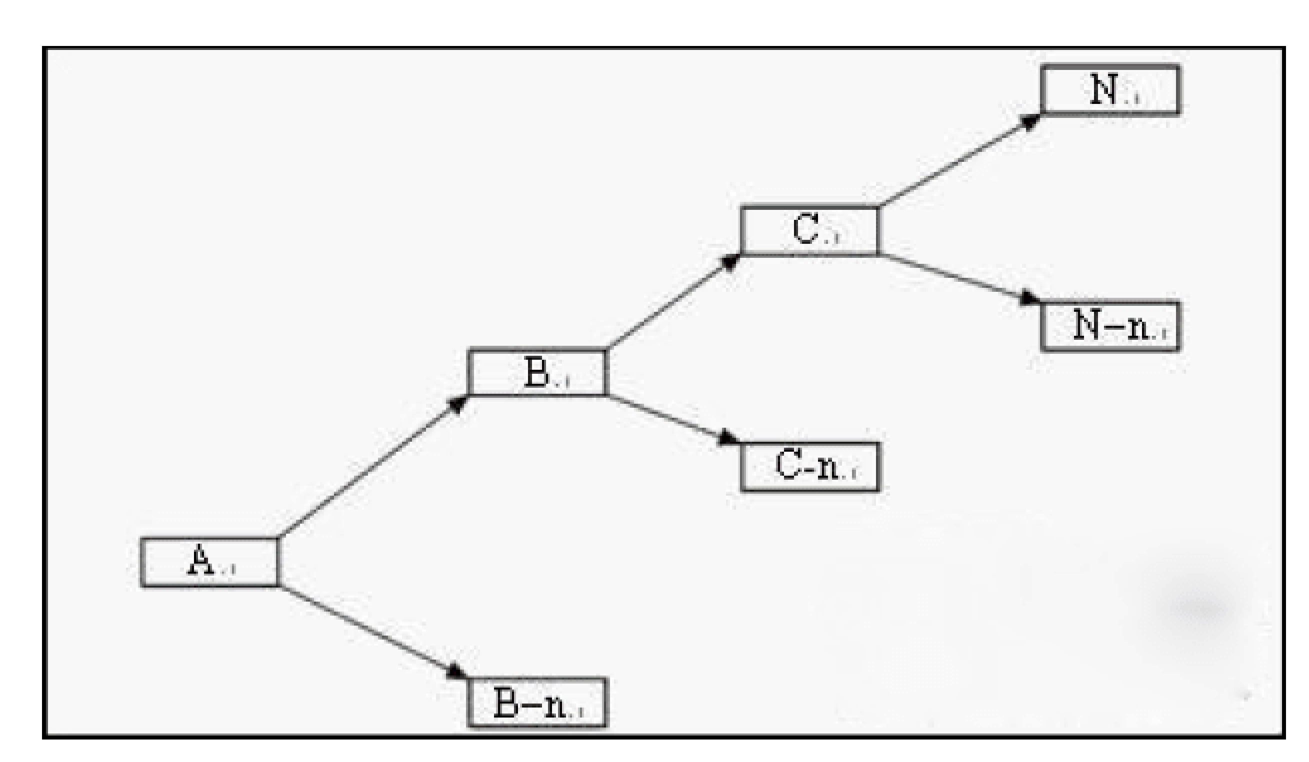


图 8-7　链接伙伴间的链接关系示意图

（7）链接评估

链接评估实际上就是了解潜在链接伙伴的情况，主要是查看潜在链接伙伴的 PR 值、页面包含数及交换链接页面上的导出链接数。

对于 PR 值，最好选择与自己网站级别相同的。如我们网站的 PR 值是 5，则最好选择一些 PR 值为 5 的链接伙伴。

查看链接伙伴网站的页面包含数主要是为了验证该网站 PR 值的真实性。某些人为了与 PR 值较高的网站交换链接，通过各种欺骗的方法，“提高”其网站在 Google 工具栏上的 PR 值。如一个网站的页面包含数为 10，而 PR 值却是 5，这明显是虚假的。

交换链接是最常见的网站推广方式之一，通过交换链接不仅可以吸引更多的用户访问网站，还能有效增加网站的外部链接数，提高网站的权重及相关性。由于交换链接是建立在平等的基础上，所以如果你的网站质量较差，则高质量的网站就会拒绝你的交换链接请求。

尽管，交换链接的主动权掌握在网站所有者手里。但是，要与外部网站建立链接关系，除了要花费时间寻找潜在链接伙伴外，对于每一个潜在链接伙伴，我们还必须对其进行评估，以及与其管理员进行联系、谈判等。

3. 链接诱饵

链接诱饵就是指通过内容或资源去吸引外部导入链接的行为，常用的链接诱饵包括软文、网络广告及共享软件。

（1）软文

软文指在报纸、杂志或网络等宣传载体上刊登可以提升企业品牌形象和知名度，或促进

产品销售的宣传性、阐释性文章。如特定的新闻报道、技术文章、付费短文广告、案例分析等。

在使用网络软文作为链接诱饵时，通常会在软文中加上指向目标页面的链接。这样，软文被广泛转载时，软文中的链接也会被引用，从而产生链接诱饵的作用。

然而，并不是所有的软文都能得到广泛的转载，某些软文一开始就说某某网站是做什么或者提供什么服务，这样的枪稿，你会转载吗？如何才能实现文章被广泛转载，并最终达到“病毒式”传播效果呢？常用方法包括利用热点话题、争议性事件或具有一定专业水平的文章等。

以下是笔者的一个小小的经历。

①热点话题

背景：2004 年初，Google 改进算法，导致大批网站在众多关键字搜索结果中的排名下降或者消失。

经过：当时国外就有 Seoer 发表了一篇名为“Google 基于 Hilltop 算法的更新”一文。在该文发表没多久，笔者就将其翻译成中文，并在网站上发布。

结果：该文章被疯狂转载，而且大部分都是同类站点，其中不乏高质量网站。

②争议性事件

背景：2007 年，“Google Adsense 无效点击”、“封杀账号”事件层出不穷。

经过：国内某站长以“Google Adsense 无效点击”为由起诉 Google。

结果：该事件被成千上万的同行网站、社区甚至媒体转载报道，“原告”网站被曝光无数次。

除了热点话题或者争议性事件的软文外，撰写具有专业水平或者对某个事件、现象，发表自己独到见解的文章，同样会被广泛转载，从而发挥文章的链接诱饵作用。

相对于其他的链接诱饵，软文具有成本低、单向性、传播性强、停留时间久等优点。但是，随着软文作为链接诱饵被广泛使用，转载者为了把链接据为己有，不但修改或删除软文里的链接，就连原作者名称也会被篡改。因此，使用软文作为链接诱饵所产生的作用已经非常有限。

在使用链接诱饵的方式增加外部导入链接时，需要注意链接的“锚文本”及“链接路径”。

（1）锚文本

在链接诱饵中，应优先使用文本作为链接对象，且以目标页面的主题名称作为锚文本，这样才能进一步发挥链接诱饵的作用。如目标页面的主题是“网站优化”，则我们在“链接诱饵”中最好使用关键字“网站优化”作为锚文本。

（2）链接路径

链接路径分为绝对路径和相对路径。

绝对路径是指文件在网站中存放的真实路径。如：http://www.123pet.com.cn/action-

news.html 就是指文件 action-news.html 存放在网站 www.123pet.com.cn 的根目录下。

相对路径是指文件相对于其他文件的存储位置（格式为“/目录名称/文件名”）。在网站内部，我们常使用相对路径引用相应的文件。如果对某一文件的引用是使用相对路径，则该文件的 URL 会随着服务器地址的变化而变化。

如果链接地址使用相对路径（如/seo/wangzhanyouhua.htm），则该链接所在的内容被转载后就会出现下面这样的情况。

- 如果内容被网站www.123pet.com.cn转载，则链接地址就是：www.123pet.com.cn/seo/wangzhanyouhua.htm。
- 如果内容被网站www.10flash.net转载，则链接地址就是 www.10flash.net/seo/wangzhanyouhua.htm。

也就是说，如果链接地址使用相对路径，那么，即使内容被转载，链接也只属于转载网站，而与原创内容的页面无关，也就发挥不了链接诱饵的作用。因此，链接诱饵中的链接地址必须使用绝对路径。

（2）网络广告

利用网络广告可以在短期内迅速提高网站的外部链接数量，从而提高页面的权重及相关性。常见的网络广告载体包括文本、图片及多媒体三种。

①文本

通过文本广告不但可以有效提高目标页面的权重，还能提高目标页面的相关性。因此，如果投放广告的目的是为了提高外部链接数量，则应优先采用文本形式的广告。而且，文本广告的链接必须直接指向目标页面。

直接指向就是指链接不经过任何形式的重定向即可到达目标页面。在 HTML 代码中形式为“关键字”，例如：

```
<a href="http://www.seochat.org">网站优化</a>。
```

如果链接要经过重定向后才能到达目标页面，则该目标页面将继承不到任何权重，例如：

- 以 SWF 文件实现的重定向，只需要在 SWF 文件里写上实现重定向的代码即可。如：网站优化，该代码的作用是，当用户点击链接后，先下载并执行文件 abc.swf，再按 abc.swf 制定的规则转向相应的目标页面。
- 以程序方式实现的重定向，常用于对广告点击数的统计。例如：网站优化表示当用户点击该链接后，先下载并执行 linkpage.php，再按照 linkpage.php 制定的规则转向相应的目标页面。

利用网络广告除了可以在短期内增加足够多高质量、单向性的反向链接外，还能自由选择链接在源页面中的位置、锚文本及样式等（如加粗、字体颜色等）。但是，网站广告存在的时间是有限的，目标页面继承源页面的权重会随着广告投放的结束而逐渐消失。

而且，目前关键字链接广告已经被搜索引擎纳入链接作弊的范围，搜索引擎也正在加强打击这种购买链接的行为，一旦被扣上“购买链接”的帽子，降权封杀在所难免。

购买链接的行为具有以下特征。

- 链接单向性。
- 使用与目标页面主题相关的关键字作为锚文本。
- 出售链接的网站中大部分的页面上都存在该链接。

如小游戏网站 www.10flash.net 在网站 A 上购买链接，则网站 A 的大部分页面上都存在以“小游戏”作为锚文本的指向 www.10flash.net 的链接，且 www.10flash.net 上不存在任何指向网站 A 的链接。

②图片及多媒体

图片及多媒体形式的广告在提高目标页面相关性方面所起的作用是极其有限的，特别是多媒体形式广告，其链接指向的目标页面甚至继承不到源页面的权重。

（3）共享软件

共享软件也是一种链接诱饵。按照软件的使用方式，共享软件可以分成为在线版与安装版，相比之下在线版共享软件更能发挥链接诱饵的作用。

①在线版共享软件

通常，在线版共享软件会以代码的形式提供给其他网站使用。要获取该软件的使用权，只须把相应的代码嵌套到页面上即可，只要该软件被广泛传播即可达到链接诱饵的作用。如开发者把其网站的链接放到共享软件的界面上，这样，其他网站在获取该软件的使用权时，该链接就会被引用，从而成为了导入链接。

②安装版共享软件

严格地说，安装版共享软件是发挥不了链接诱饵作用的，因为搜索引擎并不能识别软件界面上的链接信息。但安装版共享软件被其他网站提供下载时，会在软件介绍信息中注明软件的官方网址。这样，如果软件被越多的网站提供下载，则官方网站就可以得到越多的导入链接。

一款优秀的安装版共享软件，其达到的链接诱饵效果并不比软文差。如大家熟悉的 Windows 优化大师等。但是，软件开发成本较高，且安装版共享软件本身并不是一个链接诱饵，如果提供下载的站点不加上官方网站的链接，则官方网站就得不到导入链接，也就起不到链接诱饵的作用。

8.4.3 外部链接与网页的相关性

SEO 网站链接应该来自相关网页这一点越来越重要。一般认为，来自相关网页的链接有

更高的权重（不是 PR）。所以在寻找潜在的反向链接时，应该在与你行业相关的网站中花时间。

网站管理员会倾向于在所有反向链接锚文字里使用重要关键词。实际上这是带有一定危险的，因为显得很不自然。自发性的、自然的反向链接必然会使用各种各样的链接文字，而不可能是整齐化的。

链接应该在网页不同的位置。反向链接最好不要都来自于网页底部的友情链接区，而应该在各个网页的不同地方。可能是正文，可能是网站导航，也可能是在底部。

逐步自然增长不要在短时间内增加成百上千个反向链接，这种不自然的快速增长，是网站进入 Sandbox（沙盒效应）的重要原因之一，更严重的有可能被惩罚。

反向链接应该来自不同 PR 的网页，当然来自高 PR 值网页的反向链接很好，但如果你所有的反向链接都来自于 PR4、PR5 以上的网页，这恐怕也不自然。

来自好的邻居互相链接的网站会构成一个类似社区的网络，这个网络里的网站质量和相关性都是互相影响的。如果你的网站中很多的反向链接是来自色情网站、赌博网站、卖伟哥的网站，这不是什么好兆头。

链接应该来自不同的 IP 地址同样也是为了使反向链接趋向自然。如果你反向链接中的一大部分都局限于几个 IP 地址或几个服务器，那么试图操纵搜索引擎的痕迹就比较重。

应该有来自新旧网站的链接，这也是为了使反向链接更自然。

虽然友情链接的重要性越来越低，但友情链接毕竟是一件很自然的事情，只要不过度，不纯粹为了增加 PR 值，还是值得做一些的。

链接到内容网页千万不要把所有链接都指向你的主页，而应该把一部分链接分散到其他内容网页，同时链接文字也要有相应的变化。博客就是一个很好的例子，很多博客之间在互相引用的时候，都不会链接到主页，而是链接到文章页。

避免购买链接应该使你的反向链接看起来不像是买的、赞助的或广告，而是像自然正常的链接。链接向你的网页是不是还链接向其他的不好的网页。

8.4.4　外部链接与网页权重的继承

链接反映的是页面间的信任关系，搜索引擎根据页面的导入链接数来统计页面的得票数，从而计算出每个页面的链接权重。页面得到的投票越多，从一定程度上反映该页面的重要性越高，链接权重越大。如一个站点首页得到 100 个导入链接（包括内部及外部），而另一个站点首页得到 1000 个导入链接（包括内部及外部），假如这两个网站得到的导入链接质量是相等的，那么后者的权重明显大于前者。

通常，每个源页面中都会存在多个导出链接（包括网站内部及外部），这就涉及源页面的权重分配，或者目标页面对源页面权重继承的问题。一般情况下，决定目标页面继承源页面权重的主要因素包括：

- 目标页面的链接在源页面中的位置。

- 目标页面的链接在源页面中存在的时间。
- 源页面中导出链接的数量。

1. 链接位置

链接在源页面中出现的位置会在一定程度上影响目标页面对源页面权重的继承。如指向同样页面的链接，如果出现在源页面的左上方，其继承到的权重会远大于右下方。在规划页面的链接分布时，我们应遵循页面重要区域的分布规律："左上→右上→左→右→左下→右下"，即把指向相对重要页面的链接放在源页面的左上方。这样，该链接指向的目标页面就可以继承到更多的权重。

2. 链接存在时间

由于页面内容是经常更新的，页面中链接的更替也是正常的现象。如果指向某页面的链接在搜索引擎更新源页面前就被替换掉，则该链接指向的目标页面就继承不到源页面的权重，即对于搜索引擎来说，指向该目标页面的链接没有在源页面中出现过。

相反，如果源页面被搜索引擎更新 N 次后，指向某目标页面的链接依然存在，则该目标页面就可以继承到更多的权重，即链接在源页面中存在的时间越长，其指向的目标页面继承到的权重就越多。

假如笔者跟站点交换了链接，把他们的链接放到笔者的首页，但在搜索引擎更新前笔者就把这个链接去掉，那么这个链接就等于没有出现过，也就没有任何意义了。

3. 导出链接数量

忽略链接在页面中出现的位置及存在时间，目标页面对源页面权重的继承是以平均的方式进行的，即源页面上导出链接数越多，目标页面能继承到的权重就会越少。如站点的首页中有 N（N>=0）个导出链接，则这 N 个导出链接指向的目标页面将以平均的方式继承源页面的权重，即每个目标页面将继承到站点首页 1/N 的权重；当 N 越大的时候，目标页面能继承到的权重就越少。

此外，搜索引擎对每个页面的导出链接数量是有一定限制的，我们不能在同一页面上堆放过多的导出链接（特别是外部导出链接）。否则，不但目标页面能继承到的权重会少之又少；对于页面本身而言，也可能会由于过多的导出链接而被搜索引擎判为垃圾链接页面。

8.5 小结

链接是影响页面权重及相关性最重要、也是最难掌控的因素。本章从内部链接与外部链接两个方面，介绍链接的重要性及对页面的影响。此外，还介绍了几种有效增加外部链接的方法。希望读者能从本章得到一些启发，进一步提高网站在搜索引擎中的表现。

第 9 章　关于优化误区与作弊

搜索引擎优化的作弊是针对于搜索引擎算法的不全面而采取对应的欺骗性手段，提升了页面权重及相关性的行为，按照作弊的范围，可以划分为内部作弊以及外部作弊。本章将重点介绍内部作弊和外部作弊的基础知识，以免在实际的优化工作中重蹈覆辙。

9.1 搜索优化作弊方法

对于搜索引擎优化操作还应该掌握一个度的问题，如果操作过度，可能会被搜索引擎认为你在作弊，这样非但不会有好的排名，还会受到搜索引擎的惩罚。下面来看看哪些 SEO 优化操作会被认为是搜索引擎优化作弊的方法。

1. 隐藏文本、隐藏链接

大多是指网页专为搜索引擎所设计，但是普通访问者没法看到文本内容及链接。在形形色色的隐藏技术中，最常见的就是把文本或链接文字的字体颜色改为与背景色相同或十分接似。

（1）隐藏文本内容（Invisable、hidden text）

通常是在不影响网站美观的前提下，通过蕴含着大量的关键字，网页提高了关键字相应得分，从而直接达到了搜索引擎的排名目的。

（2）隐藏链接（Invisable、hidden links）

通常是在不影响网站美观的前提下，通过了在其他页面的渠道添加了指向目标优化页的隐形链接，通过提高链接得分从而直接改善了搜索引擎排名进度。

几乎搜索引擎都能检测隐藏技术，并查为作弊。从而包含着隐藏文本的网页，面临着被搜索引擎降低排名甚至删除列表的惩罚。虽然在 Google 上不乏使用隐藏技术而侥幸逃脱的网站，但多数人还是认为不值得冒这个险。其实通过添加可视文本内容并保证一定的关键词密度也可达到相同的优化效果。

2. 网页与 Google 描述不符

通常是先向搜索引擎提交一个网站，等这个网站正式被收录后再以其他页面替换该网站。“诱饵行为”（Bait-&-Switch）就在此类偷梁换柱之列，创建一个优化页和一个普通页，然后把优化页提交给搜索引擎，当优化页被搜索引擎收录后再以普通页取而代之。

3. 误导性或重复性关键词

（1）误导性关键词（Misleading Words）

在页面中使用了与该网页毫无干系的误导性关键字来吸引查询该主题的访问者及流量。这种做法严重影响了搜索引擎所提供结果的相关性和客观性，为搜索引擎所深恶痛绝。

（2）重复性关键词（Repeated Words）

这种作弊技术堪称为“关键词的堆砌欺骗”（Keyword Stuffing），它利用了搜索引擎对网页正文及标题中出现的关键字高度来关注对关键词进行不合理的（过度）重复。类似的其他做法还包括在 HTML 元标识中大量堆砌关键字或使用多个关键字元标识来提高关键词的相关性。这种技术很容易被搜索引擎察觉并受到相应惩罚。

4. 隐形页面（Cloaked Page）

针对于实际的访问者或搜索引擎任何一方屏蔽了真实网站内容，以供搜索引擎提供非真实的搜索引擎友好的内容提升排名。

5. 欺骗性重定向（Deceptive redirects）

把用户访问的第一个页面（着陆页）快速重定向至一个内容完全不同的页面。

“鬼域”（Shadow Domain），这是最为常见的重定向技术，通过欺骗性重定向使用户访问另外一个网站或页面。一般利用 HTML 刷新标识（Meta Refresh）来实现。大多搜索引擎都不会索引使用该标识的页面。如果你使用刷新标识来检测浏览器或分辨率，那么就请使用 Javascript，且尽可能延长重定向时间。如果一定要使用重定向功能，那么请务必确保刷新周期不少于 10 秒钟。

还有一种状况就是当用户打开一个网站，此网站自称其网站已移至新域名下，并请用户点击新域名链直接进入网站。但当用户进去后才发现，这个链接是一个“会员”链接。这也属欺骗性重定向行为。

6. 门页（Doorway Page）

也叫“Bridge、Portal、Jump，Entry Page”。是单独为某些关键字特别制作的页面，也是为搜索引擎设计的，目的就是提高特定关键词在搜索引擎中的排名次数所设计的富含目标关键词的域名，且重定向至另一域名的真实网站。搜索引擎的 Spiders 往往会忽略那些自动重

定向到其他页的页面的检索。

7. 复制的站点或网页

最经常看见的当属镜像站点（Mirror Sites）。通过了复制网站及页面的内容并划分为不同的域名和服务器，以此欺骗搜索引擎对同一站点或同一页面进行多次索引。现在大多数搜索引擎都提供有能够检测镜像站点的适当的过滤系统，一旦发觉镜像站点，则源站点和镜像站点都会被从索引数据库中删除。

8. 链接（Link Spamming）

典型的作弊链接技术包有链接工厂（link farms），大宗链接交换程序（Bulk Link Exchange Programs）和交叉链接（Cross Link）。

链接工厂（亦称大量链接机制）是指由许多网页交叉的链接而构成的一个网络系统。这些网页可能来自同一个域或多个不同的域，甚至可能来自不同的服务器。一个站点加入这样一个链接工厂后，一方面它可得到来自该系统中所有网页的链接，同时作为交换它需要“奉献”自己的链接，藉此方法来提升链接得分，从而达到干预链接得分的目的。如今搜索引擎发现它们已然不费吹灰之力。只要发现，不管属有意还是无意，链接工厂中的所有站点都难逃干系。

9. 其他

如日志欺骗行为。了解到一些页面等级比较高的站点进行多个虚假点击以求名列，从这些站点的最高引用者日志中，获得它们的导入链接。Google 已然意识到这种行为并将对利用这种途径获得导入链接的站点采取相应措施。

10. 门域（Doorway Domain）

单独只为提高特定的关键词在搜索引擎中的排名所设计的包含目标关键词的域名，然后在重定向至其他域名的主页。由于搜索引擎一般忽略自动重定向至其他页的页面检索，所以不提倡使用这种技术。

9.2 隐藏文本信息

隐藏文本的方式主要有与背景颜色相同、字体超小、display:none 和绝对定位。

1. 与背景颜色相同

假如背景是白色的，我们把文本也重置为白色，则在正常情况下是看不到的，在一些论坛里面经常用这种手法来发外链，我们可以用 CTRL+A 组合键来进行辨别。有些网页在用户

复制并粘贴内容后会出现版权信息，此时，只要用户在复制前用 CTRL+A 组合键也能看到，当然这只针对版权信息与背景颜色相同的页面。

2. 字体超小

我们在用 DW 的时候，正常情况下，最小的字体只能设置为 9px，当然也可以指定为 1px，那么就非常小了。

3. Display:none

许多网页为了实现滑动门的效果，通常用此方法，如 163 的主页，就是如此。

4. 绝对定位

一般用 CSS 来实现，先定义宽和高，然后再用 position: absolute 来进行绝对定位，如 left: -999px 等。

上述几种隐蔽文本信息的方法是我们常见的手法，也是搜索引擎最容易识别的手法。

而假如背景使用花色图片，颜色编码差异，JS 控制显示样式（友情链接欺骗常用的手法）这对于搜索引擎来说是很难识别的，也就是说这种手法相对来说高级一些，其存活率稍高一些。

例子 1：

```
<style type="text/css">
#n{display:none;}
</style>
</head>
<body>
<div><a href="http://www.123pet.com.cn">宠物网</a></div>
<div id="n">宠物网</div>
```

分析：通过 display:none 的设置把后面“宠物网”的这段文字给隐藏了，前面没有定义过的 div 则正常显示。可复制进文档，进行本地测试。

另一种隐藏文本的方法为：visibility: hidden 可见性为隐藏。

例子 2：

```
<style type="text/css">
#n{visibility: hidden;}
</style>
</head>
<body>
<div><a href="http://www.123pet.com.cn">宠物网</a></div>
<div id="n">宠物论坛客</div>
```

分析：隐藏的内容会占据它所应该占据的物理空间。

另一种隐藏文本的方法为：overflow:hidden 设置。

例子 3：

```
<style type="text/css">
#n{overflow:hidden;
width:0px;
height:0px;}
</style>
</head>
<body>
<div><a href="http://www.123pet.com.cn">宠物网</a></div>
<div id="n">宠物论坛</div>
```

分析：这段代码比较简单，这里就不详细说明了，读者可以在浏览器上运行这个网页，看看效果。

9.3 网站镜像

镜像的网站是将一个完全一样的站点放到不同服务器里，分别有自己的 URL，在这些服务器上互为镜像网站。它和主站并没有太大差别，或者可算是为主站作的后备措施。有了镜像网站的好处是：如果不能对主站作正常访问（如某个服务器坏死或出了意外），但仍能通过其他服务器正常浏览。相对来说主站在速度等各方面比镜像站点略胜一筹。

镜像就是做一个与你的网站相同的站。通常是个人网站的站长为自己的站做一个拷贝，也就是说，一个站由于订单或其他原因访问不到时，大家可以去另一个一模一样的站中浏览，这个站还可以起到分流，减少服务器压力的作用。 不过，这些都是由于个人网站的服务器不能接受太多的访问量而采取的办法。商业网站一般不用这样的办法，因为用户要记两个域名，内容要上传两次，要随时保持两个地方一致，内容一多就非常麻烦。

9.3.1 网站同步镜像

下面来看如何做网站同步镜像。

1. 介绍

目前的网站跟随着访问流量的增加，单个的服务器已经无法承担巨大的访问量。目前能有便捷的方式来解决这个问题吗？答案是“有”。如建立服务器群，进行均衡负载。

但是假如要解决如电信网通这样的互访问题，这个解决方案就无能为力了。

要解决这个问题最便捷的方法就是建立镜像网站。由访问者亲自选择合适自己网络速度

最快的网站，这样即可以解决线路问题，又可以解决访问量问题。

2. 网站同步的数据分类

网站数据可以分为以下两类。

- 一类是文件，如 HTML、ASP、PHP 等网页文件，或者 RAR、ZIP、RM，AVI 等可下载文件。要实现他们的同步很简单，用 FTP 同步软件即可。
- 一类是数据库数据文件，如 MySQL、SQL、Server 等。

数据库同步的方法也有很多种，其中最简单的就是同步数据库目录。这在后面章节会作详细讲解。

3. 网站文件的同步

在这里用到的主要工具就是 FTP，网站文件同步分两种情况：一种是本地到远程，一种是远程到远程。第一种不用说了，第二种远程到远程即 FXP，支持它的软件有很多，但是真正适合多网站同步镜像的却不多。

下面主要介绍两款网络同步软件。

（1）首先要推荐的是国产的 FTP 软件“网络传神”，功能非常强大，特别是在网站的同步镜像方面。可惜的是，这款经典的软件已经不再更新了，该软件的最后更新时间为 2003 年 3 月，最后一个版本为 3.12，即便如此它还是非常好用的，下面是一段官方的简介。

网络传神吸收了 Cuteftp 和 UpdataNow 的全部功能，并且增加了其他软件没有的多项功能：支持网站互传，支持网站同步（UPDATA NOW），支持后台上传（多线程上传多个文件），可同时打开多个站点、多站点计划上传功能，支持镜像站点，支持宏操作支持计划操作，支持文件高级比较上传，支持目录隐藏过滤（为用 ForntPage 作主页的朋友带来福音），服务器自动识别功能，资源管理器浏览方式，可以自定义命令，支持 RFC959 标准具有更好的稳定性及完备的信息返回机制及错误监控机制完整的中文帮助。

（2）由 ReGet 同一开发公司制作的专用于网站同步的软件“WebSynchronizer”，用这款软件，访问者会体验到网站同步的方便快捷、简单容易。其最新版本为 1.3.62，网上能找到的最后版本是 1.1 版，下面是一段简介。

档案同步化工具 WebSynchronizer，由知名续传软件 ReGet 之软件出版公司所推出，是网站同步化、档案镜像、档案备份的绝佳工具，可以执行以下主要工作：

- 本机资料夹及远程资料夹的同步化。
- 两台远程计算机中的资料夹同步化。
- 两个本机资料之同步化。

其他还有一些软件如同步快梭（AutoSyncFTP），也能实现简单的网站同步。不过，这款软件非常不稳定，而且 2001 年就已经停止开发。所以，不用考虑了。还有 SiteMirro 这款软件，由于网上找不到可以用的版本，所以没有办法测试。

9.3.2　如何做同步镜像

母站设置最重要的地方就是“比较目录”（网站同步）。

1. 拖动传送

就是手动拖动传送时候的传送方式，非拖动方式不受影响，因为母站的更新基本上是从本地目录进行的，所以建议这里选择“只上传不同的文件”，镜像站点这里可以忽略。

2. 磁盘比较设置

这个功能非常有用，可以通过文件长度和文件日期进行文件差异对比。还可以设置本地目录过滤和服务器目录过滤，非常灵活。

添加完母站和镜像站之后，就可以设置同步了。如果是添加镜像站，则会跳出母站选择窗口。设置好多个镜像站点之后，返回主界面就可以进行网站同步了。

设置时要注意，一般不要改动镜像站点前面的“对钩”标志，因为镜像站点同步的原理是“根据主站点的同步信息更新镜像站点”，如果镜象站点没有更新那么必须将该镜像站点单独同步。通常应该选择“选项窗口”→“FTP 客户端”→“计划”中的“最大连接个数”和“最大线程个数”，一般不要改动保持为“1”。

如果同步时某个站点出错同步没有完全完成，只须按“计划”窗口中相应的按钮单独执行某个站点的传送即可。

3. 点击设置

进入网络传神的选项设置窗口，选中 FTP 客户端中的定时同步。

- 定时下载：在指定的时段内下载文件。
- 定时上传：在指定的时段内用网站同步的方法进行上传。
- 开始时间：如果要使用定时方式必须设置此项，在设置的时间中会自动开始所有设置为计划的站点。
- 停止时间：如果设置此项则当系统时间大于停止时间时自动停止所有设置为计划的站点。

注意，定时上传没有“停止时间”选项，网站同步完成后会自动停止，同时网站同步同一时间只能运行一次。

网络传神还有很多为便于同步镜像设置的功能，大家在实际使用中去体会吧。

4. Web Synchronizer 网站同步镜像使用傻瓜指南

Web Synchronizer 功能非常强大，不单单是网站同步，还可以进行本地目录同步，本地到远程同步，功能方面远胜过网络传神。不过，国内没有汉化版，对于一些对英文不太好的朋友用起来会不太舒服，希望 CCF 里的汉化高手能将它中文化。

开始运行程序，“老外”的软件就是智能，打开该软件就会有一个同步向导跳出来。

在这里，可以看到该软件只支持两个目录间的同步。不过，我们可以创建多个项目来解决这个问题。具体的将在后面介绍，这里我们来看看如何建立两个远程目录的同步。

选择第二项，进入下一个选择框。

要进行网站同步，此时应该选择第一项，再进入下一步，再选择同步的第一个服务器，建议在 servers list 中，先编辑好需要同步的几个 FTP。


提示

同步方式设置说明

第一项的意思是上传下载所有改变的文件。

第二项的意思是不改变文件夹内容，只下载新的文件。

点击下一步进行第二个服务器设置，方法和第一个一样。

设置完成后，你还可以修改任务名称（个人觉得这点就是该软件非常优秀的地方，以任务的方式进行管理，非常方便），此时就可以看到他的主界面了。

有朋友会问，如果要进行两个以上的网站同步应该怎么办?

很简单，再多建几个任务就可以了，要镜像多少网站都没有问题。

这个软件的网站同步镜像可以非常灵活。如可以本地→远程，然后再远程→远程，或者本地→远程 1、本地→远程 2、远程 1→远程 2，任何情形下的同步都可以完成。

9.4 门页

管理员建立一个门户页面只有一个目的，使某些关键字或短语在某些搜索引擎上获得更好的排名。假设你开了一家 Pizza 餐厅，肯定有很多人来你的餐厅，但不是所以的人都要同一口味的 Pizza。搜索引擎的工作原理也是一样，他们不是都喜欢同一个页面的。有些时候你的网站在某个搜索引擎获得了较好的排名，但是在其他的搜索引擎中却找不到。所以是时候去建立一些门户页面去适应多个搜索引擎的“口味”。如果搜索引擎发现你的门户页面适合他的口味，你网站排名将会提高。

1. 垃圾与门户页面

在过去，门户页面只是由几个关键字或者几幅图片组成，没有实际、有价值的内容。然后直接链到主站。人们喜欢使用一些软件自动生成多个门户页面，然后自动向众多搜索引擎提交，结果排名却反而更后了。那么门户页面对于网站排名真的起到了积极的作用了吗？有些人还在不断的争论，这样的门户页面对某些搜索引擎来说仍然是有用的。低质量的门户页面没有实际的内容，则会被搜索引擎认为是垃圾。

2. 对搜索引擎友好的门户页面的组建

如果使用得当，是不会被搜索引擎误认为是垃圾。如果能建立高质量的门户页面，这对于提高网站排名会起到一个非常积极的作用。如果要用门户页面来提高你网站排名，你的门户页面必须给访问者提供足够多的信息、图片和链接，正如你的主站一样。一个好的门户页面应该像正式的页面一样，只不过是根据特定的搜索引擎而设计。以下几点要切记。

（1）没有必要对每一个搜索引擎建立相对的门户页面。把你的注意力放在那些对你网站没有好感的搜索引擎上（当然，如果所有的搜索引擎都不喜欢，那么你就没有必要去建立门户页面，此时做好你的主站就行）。

（2）不要向主要的目录提交你的门户页面，因为这些目录只允许收录你的主站页面。否则，你将会受到惩罚。

（3）在你建立门户页面前，必须对所针对的搜擎引擎有比较深的了解。

（4）有效的门户页面，只会用到 1~2 个关键字。如果用得太多，则会有相反的效果。把注意力尽量集中在你想优化的几个关键字上。

（5）把门户页面建得尽量普通（即与正常的页面一样，不要让人一看就知道这是一个门户页面）。

3. 伪装

伪装是指根据用户返回不同页面的行为，是门页最常用的跳转方式之一。如向搜索引擎返回经过特别优化的页面，而对普通用户则返回正常的页面。不管在什么情况下，伪装都是欺骗搜索引擎的违规行为，因此会受到搜索引擎的严惩。

9.5 重定向

重定向就是通过各种方法将不同的网络请求重新定个方向转到其他位置（如网页重定向、域名的重定向，路由选择的变化也是对数据报文经由路径的一种重定向）。

常用的重定向方式有：301 重定向、302 重定向、meta fresh 重定向与 JavaScript 重定向。其中，301 重定向和 302 重定向在 4.7 节中已经介绍过，这里就不再赘述，本节主要介绍 meta fresh 重定向和 JavaScript 重定向。

9.5.1 meta fresh 重定向

meta fresh 重定向是指通过设置<meta>标签的 http-equiv 属性值及内容来实现的重定向。如在页面头部加上代码：<meta heep-equiv="refresh" content="3;url=http://www.123pet.com.cn">，则打当前页面自动跳转至 www.123pet.com.cn。

在 meta fresh 重定向中，如果设定的停留时间果断，如果少于 1 秒，则会被搜索引擎认为是门页。为了避免门页嫌疑，在使用 meta 跳转时，通常会把停留的时间设定在 3 秒以上。

9.5.2 JavaScript 重定向

本例定义了两个 JavaScript 函数，功能都是重定向到首页，但是第一个函数采用的是直接给 href 赋值的方式，第二个使用的是 replace 方法。具体请看下面的代码：

```
<script type="text/javascript">
function goBack () {
location.href = "http://www.123pet.com.cn";
    }
function goBackRep () {
location.replace ("http://www.123pet.com.cn") ;
}
</script>
HTML 代码
<p>
点击后面的按钮就可以使得页面转向首页，
可以点击浏览器的后退按钮返回本页。
使用的方法是直接给 location.href 赋值。
</p>
<button onclick="goBack () ; return false;" value="重定向到首页">
重定向到首页</button>
<p>
点击后面的按钮同样会跳转到首页，但是，
由于是使用 location.replace 的方法，所以不能。
通过浏览器的后退按钮返回本页。
</p>
<button onclick="goBackRep () ; return false;" value="重定向到首页（不能返回）"
>
```

运行后的效果：点击后面的按钮就可以使页面转向首页，点击浏览器的后退按钮则可返回本页。使用的方法是直接给 location.href 赋值。

9.6 链接作弊

链接作弊的方法多种多样，如利用很多站长不懂导出权重的原理，通过交叉链形式来作弊，其隐蔽性高。表现：一般对方会用一个 PR 比较高的导航站给你网站做链接，然后让你的网站给他的另一个站做链接。看上去很公平，这里一定要注意，交换友链不能只看 PR

值，很有可能对方买了一个高 PR 的域名，然后改个导航站，就拿它去招摇撞骗了。下面介绍几种链接作弊的情况。

1. 通过 JS 代码输出友情链接

原理：因为搜索引擎不识别 JS 代码中的程序内容，所以如果对方的友情链接是 JS 代码输出，则完全达不到友链传递权重的效果。也就相当于，你给了对方一个单链。

表现：对方经常会截图给你看，然后让你去点击。JS 的链接点击当然是有效的，完全可以点击回我们自己的网站。

对策：一定要检查源代码，仔细看友链实现部分的代码，如果是 JS 调用，则说明这就是一个典型的友链作弊。



有些朋友可能对代码不熟悉，因为 JS 写法多种，在这里无法一一说明。不过针对这种作弊，有个好办法，直接在源代码中查找你的锚文本关键词或你的域名，是查找不到的。通过这个方法很容易识别。

2. JS 代码变种形式：onclick

原理：和第一点一致。

表现：一般链接标签的写法是 锚文本，这种作弊是利用 onclick 这个 JavaScript 里的事件函数进行跳转，如：onClick=location.href=“链接地址”。

对策：仔细查看友链代码部分的内容，看看<a>标签的写法是否使用 onclick。

3. 框架嵌套 iframe 作弊

原理：利用网页 iframe 框架嵌套新的页面，然后在新页面里显示友情链接。这样的友链它并不能让对方首页等高权重页面传递过来，实际情况是 iframe 嵌套的页面才是真正的友链页面，而这种页面基本是不收录的，就算收录，也肯定没有多少权重。

表现：链接正常，所在页面代码也正常，点击反应也完全没问题。这种作弊的破绽在于 iframe 本身。

对策：先在首页源代码里查找我们的锚文本关键词或域名，如果发现无法查找到，再查找“iframe”这个英文代码，确定友链代码是在 iframe 嵌套的页面中，就基本判断出作弊了。如：<IFRAME height=243 src="假的友链页面网址" frameBorder=0 width="101%" scrolling=no></IFRAME>。

4. nofollow 属性作弊

原理：nofollow 属性告诉搜索引擎，这个链接权重不随着链接导出，这样做的友链就没有根本的意义了，属于典型作弊之一。

表现：有两种，一种是直接在<a>标签中写上 nofollow 属性，如<A href=“友链地址”

rel=”nofollow”>友链锚文本</A>；还有一种更加隐蔽，就是通过 ID 或 JS 等手法，给<a>链接标签赋值 nofollow 属性，因为代码是分开的，而且可能藏在某个谁也不会去注意的 JS 文件中，非常不容易查询。

5. 劫持 PR 来交换友链

原理：利用 PR 劫持，然后去换取高 PR，高权重的站点友链。

表现：越是 PR 劫持的，越喜欢炫耀自己的 PR 高，交换友链的时候需注意这点。

对策：通过站长工具等去检查对方站点的 PR 值，真假 PR 很容易查询。

6. 真假页面来欺骗友链

原理：我们肉眼看到的首页和搜索引擎看到的页面并非一样，通过黑帽技术这点是完全可以做到的。

表现：就表现而言，毫无破绽。

对策：可以用工具模拟蜘蛛抓取，然后看抓取结果代码，是否有友链代码存在。主流的友链检查工具也有这个功能。

7. 高级的黑帽欺骗手法

原理：涉及黑帽技术，基本是第 6 点的升级版本，唯一的区别在于它通过蜘蛛模拟也识别不出来，而主流工具也查询不到问题。

表现：注意那些没有百度快照的，排名却非常好的网站。注意，不是快照删除，而是快照内容为空。

对策：一般利用这个技术作弊的，SEO 水平至少是黑帽而且级别不低，友链交换也会比较多，如果碰到友链非常多，而且用翻滚等形式展示的站点一定要小心，十之八九会有问。

8. 做上再删掉等手法

原理：这个没原理，真的要说个原理，只能说是人品。

表现：首先和你做上友链，然后过一段时间就删除。继续找，继续做，继续再删。也算弱智作弊手法之一，主要是利用很多站长不经常打理友链。其次是先让你做上他们的友链，接着会说他们那边的友链技术，后面会统一时间做；然后就拖着，直到你忘了这事，或是你在某次检查友链的时候结束这个行为。再者就是其他类似行为。

对策：多检查友链，每隔一段时间就检查一次，做好管理。

9.6.1 识别垃圾链接

搜索引擎识别到垃圾链接后，并不会立即对我们的网站进行惩罚，但会增加对我们网站的怀疑指数。而这种作弊手段用得太多之后，搜索引擎的惩罚就随之降临了。所以提醒大家建设外链的时候一定要避免产生垃圾链接。

1. 大量相同的主题

站长们在发外链的时候有时很懒，将内容相同的主题发布到很多论坛中。这些相同的主题对用户没有产生不同的价值，搜索引擎是非常讨厌采集非原创内容的。这样的文章大量发布，产生的链接可能会被搜索引擎判断为垃圾链接。

2. 与主题不相关的回复

有些站长在回复主题的时候并不看主题，而是自顾自地写一些与主题不相关的回复。这样的回复看似省时了。但没有与主题产生互动，是一个孤立的回复，不会对这个页面价值产生贡献。

3. 回复简单而无意义

经常看到有些站长在回复的时候就是“学习了”、“谢谢分享”、“太有道理了”这样的非常简短而毫无意义的回复。甚至回帖者压根没看清楚，其实楼主是在提问。这样的回复不仅是没有意义，甚至可以理解为恶意水帖。

4. 回复带有大量链接

有的朋友做了签名还觉得不够，在回复中还带一些链接，没有几个字的回复，但链接却不少。甚至还有一部分在回复中带隐藏链接，这些都是很容易被搜索引擎判断出来的。不仅不能给网站增加有效外链，还会增加网站的作弊风险。

5. 大量留言相似

垃圾链接常常会有一些共同点，群发软件的留言往往是内置好了的。这样的链接对页面产生贡献的概率基本没有。当有大量相似的回复链接指向同一个网站，这些链接毫无疑问会被搜索引擎判断为垃圾链接。

6. 留言链接增加过快

群发的软件不同于人工链接的主要特点就是群发的软件能够在短时间内制造大量垃圾链接。不仅是在同一个论坛增加大量的链接，也指很多论坛和博客留言在同一时间增加大量链接。

9.6.2 作弊的处罚

对于作弊被惩罚的后果是很严重的，较轻的惩罚，是系统在给网页排名时略为降低该网页的排名。较常见的惩罚，是作弊网站的某些关键词（或所有关键词）的排名降到其他任何网站之后。较严厉的惩罚，对于作弊行为和后果特别恶劣的网站，将从搜索中消失。

1. 搜索引擎作弊处罚

主流搜索引擎，如 Google、百度、Yahoo！都拥有自己的反作弊小组，专门应对搜索引

擎作弊，维护良好的搜索体验。

2. 搜索引擎作弊判别原 IP

善意原则为在确定某一网站是否有作弊嫌疑时，通过详细判断，收集证据，来辨别网站是否真有所谓的“恶盒”。

坚持自动、机器判别原则针对统一标准，由更理性和细心的机器来作出判断。

3. 搜索引擎作弊处罚方式

一旦站点的搜索引擎作弊行为被搜索引擎发现，那么客户站点将会遭到降权或者直接被搜索引擎从索引中清除。

站点，几乎没有机会被重新收录，这就是片面追求搜索引擎排名效果的可悲之处。根据网页的作弊情况，会受到不同程度的惩罚。

特别补充，含恶意代码或恶意插件等用户极端反感内容的网页，虽然不是作弊，但只要用户投诉，一经确认也会被搜索引擎拒绝收录。

9.6.3 被处罚后的补救方法

在 SEO 工作者中，有 30%的人想走一些捷径的道路，但是真正走捷径成功的人又有多少呢？在 SEO 中所谓的捷径路步步是陷阱，虽然眼看你就要到终点，但是当你要走向终点的时候你却掉进了陷阱，后悔也来不及，求救别人也帮不了你，所以说做 SEO 的时候最好是使用正常优化推广。不要为了快速地看见效果就加入黑帽方法走捷径的道路，搜索引擎就像一个法官，当你触犯它的法律的时候，等待你的就只有惩罚并且让你的网站在搜索引擎中永远的挣脱不出来，所以请 SEOer 们不要掉进捷径的陷阱里去。

所谓知己知彼才能百战百胜，所以大家要知道搜索引擎的法律，以下是搜索引擎的法律，请大家遵守。

1. 不能优化的网站

一般被搜索引擎列为可以不被优化的网站的首页都是一些信得过的门户网站和一些与搜索引擎有合作的商业性网站。而搜索引擎中的法律之一就是严格规定出了不能做优化的网站，以下就是不能优化的网站。

（1）色情（成人）内容，搜索引擎已经屏蔽了相关词的收录通道，设定了只收录搜狐、雅虎等大型门户型网站提交的新闻资讯。

（2）窃听设备和服务（电话窃听器、手机窃听器、传真拦截器、针孔摄影机），盗取别人隐私的仪器等属于违法型网站，所以搜索收录通道被关闭。

（3）赌博相关行业，国家法律规定关闭。

（4）学术活动代劳服务（“枪手”服务，毕业论文、学术论文等服务），国家法律规定关闭。

（5）药检作弊助剂（禁药去除剂和尿检添加剂），药品国家统一指定安全出厂商，所以个人的企业站是无法上排名的。

（6）毒品与吸毒用具，涉及社会治安，国家法律规定关闭。

（7）与拨号器相关的产品，百度已关闭相关产品收录信息。

（8）非法交通器材（雷达干扰器、牌照遮盖设备、交通信号灯变换器及相关产品），国家法律规定关闭。

（9）武器（如枪支、枪支组件、弹药、甩刀（弹簧折刀）、蝴蝶刀和指节铜套），涉及社会治安，国家法律规定关闭。

（10）暴力或反动网站，涉及到教育，国家法律规定关闭。

（11）黑客与破解，涉及网络安全，国家法律规定关闭。

（12）群发营销产品，如电子邮件群发软件，百度已关闭相关产品信息。

（13）烟草制品与香烟，国家法律规定，指定药草出厂商。

（14）伪造文件，伪造证件扰乱社会和谐，国家法律规定关闭。

（15）诈骗网站，国家法律规定，诈骗类网站是违法网站，不许存在，一旦发现直接关闭。

以上就是不管你怎么做都不会出现在首页的网站，如果 SEOer 们接到这样的网站请不要做，不要触犯搜索引擎的优化规则，走捷径虽然有用，但是效果只有一会的时间，搜索引擎对这些相关类的网站监控很严格的，所以以上网站就放弃做 SEO 吧。

2. 搜索引擎中优化规定

搜索引擎中的法律其二就是规定出在优化的时候不能使用的方法，而且在做推广优化的时候不能超过的范围，也就是说在搜索引擎法律中在网站优化的时候也是有法律时刻跟随着你，以下是搜索引擎中规定不可以使用的方法。

（1）欺骗搜索引擎，如果你在你的页面内或者代码中加入了一些不符合你自己网站本身现有的信息，而是一些虚拟的信息，那么搜索引擎一旦发现会被视为欺骗搜索引擎，会使你的网站在搜索引擎中被封杀。

（2）文章更新的时候，关键词不能加得太多，而且关键词间隔不能小于 8%，如果小于8%，搜索引擎会认为你的网站在恶意添加关键词使得搜索引擎被误导收取你的信息。

（3）文章更新的时候，不易添加非法信息，如果加入了非法信息搜索引擎会视你的网站为非法网站，而且使你的网站在搜索引擎中被封杀。

（4）文章不宜伪原创，如果使用伪原创工具一旦被搜索引擎发现，你站内文章被收录的信息全部会被删除，导致你的站内外链全部无效，所以说不要走捷径给自己惹来“杀身之祸”，而且文章最好是原创，就算伪原创也不要用工具，用自己的话重新来描述文章的含义也可以。

（5）友情链接交换使用黑链接交换也是一种违规现象，但是这个违规现象不是很严重，搜索引擎对黑链接的查处也不是很严格，所以现在做黑链接的人越来越多的原因就是因为搜索引擎查处的几率很低。

（6）软件群发工具，只要利用程序进行大量发帖的全部都属于黑帽技术，所以说使用黑

帽技术的人一旦被搜索引擎抓到只有死路一条。

搜索引擎优化时，只要不触犯以上 6 条法律就不会有什么问题的，也不会被搜索引擎封杀。

3. 如果被封杀需要做的是什么

如果说你的网站真的被搜索引擎封杀了，那么你需要做以下的紧急措施。

（1）每天写 3~5 篇原创软文，然后投到权重比较高的原创网站和一些搜索引擎产品的论坛和贴吧里，吸引蜘蛛爬行你的网站，如果运气好的话你的网站会复活。

（2）对一些权重较高的论坛和贴吧进行发帖，也是为了吸引蜘蛛爬行你的网站。

（3）多做权重高的网站的友情链接交换，最好是 PR 值高过自己的，如果是一些大型网站就更好了。

（4）每天向搜索引擎提交 1 次网站，运气好的话则可能会被收录。

以上方法不一定 100%好使，但是这么做了你的网站就有可能会在搜索引擎中复活。

9.7 如何举报作弊网站

一般可以举报的内容包括：隐藏文字或链接、误导性或重复字词、不符合 Google 说明的页面、被遮蔽页面、欺骗性重定向、桥页、大量重复或复制的内容、购买链接等，网站被别人复制了也可以向搜索引擎举报垃圾网站。

百度的举报地址是：http://www.baidu.com/search/web_tousu.html 如图 9-1 所示。



图 9-1 百度提交作弊网站页面流程

Google 的举报地址是：https://accounts.google.com/ServiceLogin ? service=sitemaps&

passive=1209600&continue=https://www.google.com/webmasters/tools/spamreport&followup=https://www.google.com/webmasters/tools/spamreport。

通过欺骗搜索引擎来把网页排名提高被称为搜索作弊，作弊可能使站点被搜索引擎禁止收录。有很多不同的方式来欺骗搜索引擎，了解他们是为了避免我们的站点因违反规则而遭受惩罚，同时通过举报作弊网站来维护自己的权益。

在做 SEO 优化关键字排名的时候，别人资源比你多，但是使用了过多的黑帽技术，这时候，你就应该为了中国的网民做一份有意义的事情，那就是去举报你的竞争对手采用了搜索引擎作弊的手法。

下面是 4 大搜索引擎的垃圾网页投诉中心的网址。

- 向 Google 举报垃圾搜索结果：http://www.google.com/contact/spamreport.html；向 Google 举报付费链接：https://www.google.com/webmasters/tools/paidlinks？hl=zh_CN；通过 Google 网站管理员工具向 Google 举报垃圾索引网站：https://www.google.com/webmasters/tools/spamreport？hl=zh-CN。
- 百度垃圾网页投诉中心：http://utility.baidu.com/antispam/spamreport.html。
- 中国雅虎作弊/申诉网站反馈页面：http://help.cn.yahoo.com/feedback.html？product=websearch。
- 向 Sogou 举报作弊网站：sogou-antispam@sogou.com。

9.8 小结

搜索引擎优化中，有大部分的方法都属于作弊技术，不过有很多对 SEO 不是特别了解的朋友很可能会把一些作弊方法用到日常的网站优化工作当中。我想这类站长要么是新手，要么就过度轻视了搜索引擎的智能，最终导致自己网站排名的消失和网站被封杀的结果。

第 10 章　SEO 辅助工具

SEO 辅助工具就是对搜索引擎功能优化处理的工具。用搜索引擎检验网站的搜索引擎优化状况是最直接、最有效的方式，因为搜索检索结果可以带给人们大量有价值的信息。通过搜索引擎检索结果的分析，是研究网站搜索引擎优化状况的有效方法之一，因此搜索引擎优化工具，又称为 SEO 优化工具。通过其检查网站链接数量、网站被搜索引擎收录网页数量、网站的 PR 值等。它只能在一定范围内反映出某些指标的状况。应用搜索引擎优化工具的目的，是为了检验一个网站对搜索引擎的友好程度及其可能获得的搜索排名效果。

10.1　关键字查询

SEO 的工作者在分析关键字的竞争度和相关价值的时候经常会使用一些工具来节省时间并减轻自己的工作量。常用的关键字查询工具有哪些呢？下面笔者就为大家推荐一些常用的关键字查询工具。如果我们能够用好这些工具，相信会对我们的 SEO 工作有很大帮助。

1. Google Trends

谷歌趋势（Google Trends），是谷歌公司发布的一款基于搜索日志分析的应用产品，其产品通过分析谷歌数据库中的搜索结果，告诉用户某一搜索关键字在 Google 中被搜索的频率和相关统计数据。在 Google Trends 中的每一个关键字的趋势记录图形的显示，分为搜索量和新闻引用量两部分，用户可直观地看到每一个关键字在 Google 全球的搜索量和相关新闻引用情况的变化走势，并有详细的城市、国家/地区、语言柱状图显示，如图 10-1~图 10-3 所示。



图 10-1　谷歌趋势查询页面



图 10-2　谷歌趋势查询结果分析

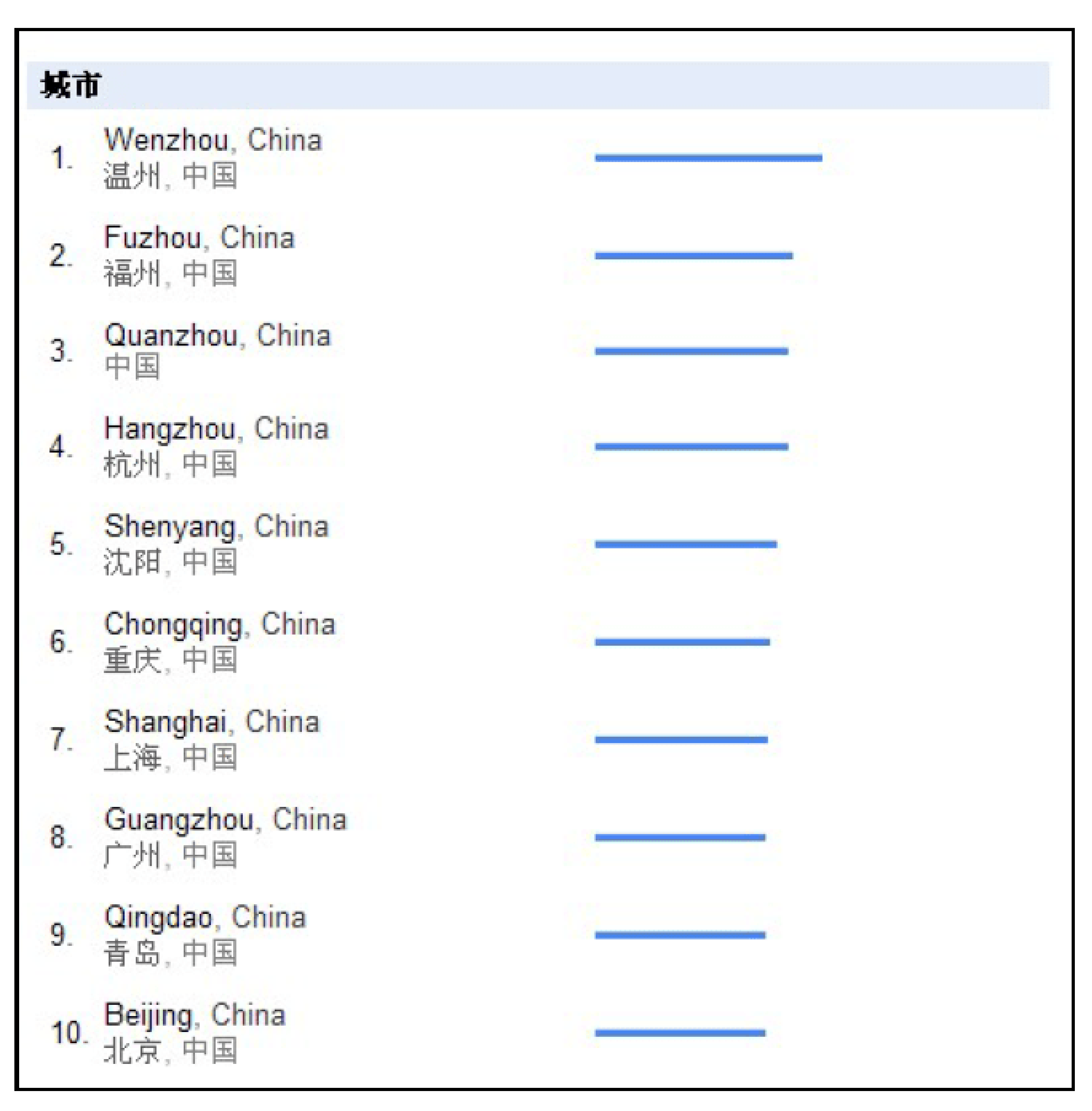


图 10-3　谷歌趋势地区分析结果

2. Google Insights Search

谷歌搜索解析（Google Insights Search），其与谷歌趋势的大部分功能都相同，只是在细节方面比 Google Trends 更加细致，所以显得更为专业。Google Insights Search 可以比较特定

区域、类别、时间范围以及搜索资源之间的搜索量，如图 10-4 所示。

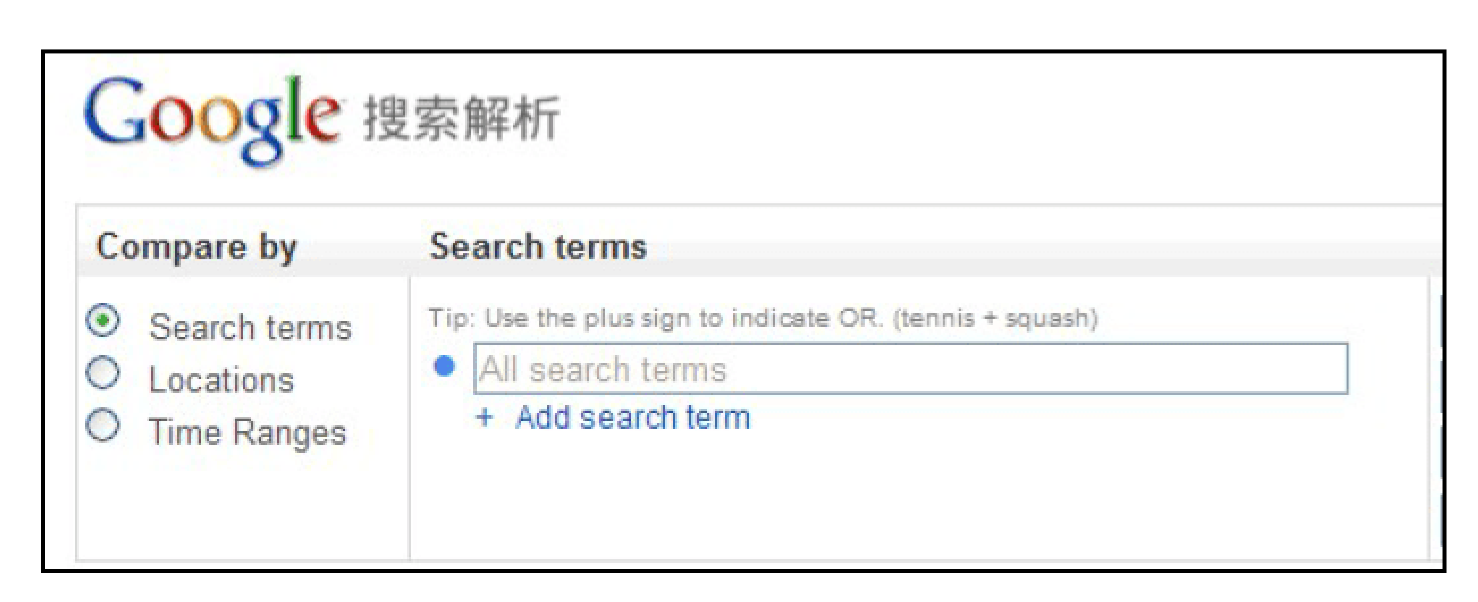


图 10-4　Google Insights Search 查询页面

3. 百度指数

百度指数是百度公司所推出的和谷歌趋势功能差不多的一款信息服务产品，不过其搜索数据是基于百度自身的，这个工具百度是以自身网页搜索和百度新闻搜索为基础的数据分析服务产品，用来分析不同关键字在过去一段时间里的“用户关注度”和“媒体关注度”，对于广大的 SEO 工作者来说，这个关键字分析工具很有用，如图 10-5 所示。



图 10-5　百度指数查询首页

4. Google AdWords 关键字工具

Google AdWords 关键字工具可以帮助使用者自己选择关键字并跟踪关键字的排名效果。可以根据用户指定的关键字自动生成最常用同义词及相关词组列表，并为用户提供关键字的点击量估算。其“流量估算工具”还可以估算指定关键字的全球每月搜索量以及本地每月搜索量，适合于挖掘信息量很大的垂直型门户网站的长尾关键词，如图 10-6 所示。



图 10-6　Google AdWords 关键字工具查询页面

10.2　Google 管理员工具

Google 网站管理员工具（Google Webmaster Tools），是目前大多站长都在使用的，它是可以免费提供网页在 Google 上展示详细报告的一个站长在线管理平台，其前身是 Google Sitemaps，以前主要的用途是让网站解决蜘蛛的抓取故障和提交 Sitemap。

Google 为了更好地和网站管理员沟通，推出了这套服务，让站长们可以在此查询网站在谷歌的索引和排名情况。可以说 Google 网站管理员工具是网站管理员与 Google 对话的平台，不管是懂 SEO 或从来没听过 SEO 这个词的网站管理员，都可以从这个最基本的 SEO 工具获得帮助。权威、免费、全面，这些特征让 Google 网站管理员工具从一上线就深受网站管理员欢迎。

10.2.1　Google 网站管理员工具的功能

笔者通过 Google 官方发布的介绍，经过整理完善之后统计出的功能基本有以下几点。

（1）从 Google 的角度查看你的网站并发现问题

查看 Google 如何抓取网站并编入索引，了解网站访问过程中的具体问题。

抓取信息：了解如何确定我们已访问过你的网站，并查看 Googlebot 上次访问的时间。

你还可以查看我们在抓取时遇到问题的网址，并了解无法抓取的原因。如此一来，你便可以修正所有问题，让我们能够为你的所有网页编制索引。

索引统计信息：了解对网站编制索引的方式以及你的哪些网页目前已编入索引中。如果发现你的网站有违规行为，我们会给你机会修正问题并申请对你的网站进行重新审核。

（2）查找你的链接并查询流量

使用新的链接报告工具查看、分类并下载有关你网站的内部和外部链接的全面数据；了解有哪些 Google 搜索查询吸引用户访问网站，并确定用户如何进入网站。

热门查询：找出提高你网站点击量的热门查询，以及你的网站出现在热门搜索结果中的位置。这可让你了解用户如何找到你的网站。

网页分析：查看你网站的热门内容以及其他网站用来链接到你网站的字词。

（3）共享有关你网站的信息

使用 Sitemap 告诉我们有关你网页的信息，哪些网页最重要以及这些网页多长时间更新一次。你也可就我们编制索引的网址外观向我们说明你的偏好。

（4）抓取速度控制

向我们提供有关你希望 Googlebot 如何抓取你网站的信息。如果你希望我们减慢或加快抓取你网站的速度，请告诉我们。

Google 管理员工具可以帮助站长解答以下疑问。

（1）谷歌是如何“看待”或理解你的网站的？

（2）他们将用什么样的关键词与你的网站内容关联在一起？

（3）你是如何内部连接网站上的各种网页的？

（4）是否有任何错误或问题会阻止谷歌抓取和索引你的整个网站内容？

（5）你的网站被搜索到的情况如何，以及用户是怎么理解你的网站的？

（6）使用哪些搜索关键词可以在搜索结果中找到你的网站？

（7）你的网站在搜索结果中出现的频率如何？

（8）用户点击你的网站的频率如何？

（9）你的网站中最成功、最受欢迎的网页是哪些？

（10）外部网站与你的网站互动情况如何？

（11）哪些外部网站可以链接到你的网站？

（12）你的网站是否会受到恶意软件的感染？

Google 网站管理员工具除了提供一些有见地的数据来回答上述的问题，还能利用 Sitemap（上传一个 XML 站点地图，能让搜索引擎更容易抓取你的网站）让你更轻松地将网站告之搜索引擎。

从 Google 网络管理员中心选择“网站管理员工具”，如果你已经拥有 Goolge 账户，不

管是 Gmail、Adsense、iGoogle 或其他产品的账号都可以直接登录。如果都没有，则注册一个即可。

10.2.2　Google 网站管理员工具的使用

Google 网站管理员工具的使用是很简单的，进入登录页面后每一步都有明确的提示，即便没有经验的人也可以很轻松地进行操作，下面给大家简单介绍一下步骤。

步骤01 使用 Google 网站管理员工具，首先必须要拥有一个谷歌账号，如果你还没有请先创建一个谷歌账号，创建完毕后，打开 Google 管理员工具的首页，输入账号和密码登录即可。

步骤02 进入后即可看到控制台，在文本框处输入你网站的域名，这里不要加 http://，然后单击“添加网站”按钮。

步骤03 接着会弹出该域名的概述页面，在左侧可以看到其他 6 个选项卡，我们需要先验证网站，才可以使用全部功能。

步骤04 点击验证网站链接，可以选择两种验证方法，一种是添加元标记；另一种是上传一个 HTML 文件，两种方法都可以。验证完毕后，即可使用全部功能，下面逐一进行讲解。

10.2.3　Google 网站管理员工具功能的应用

进入管理页面，在首页概述中可以看到整个网站的抓取和索引状况，通过这里可以看到网站是否收录正常。如果网络抓取错误，可以直接反馈 Google 蜘蛛在爬行过程中遇到的错误，以使我们可以及时地了解网站存在的问题，并及时进行修复。下面来分别介绍 Google 网站管理员工具的管理功能。

1. 基本功能介绍

很多用户在使用 Google 网站管理员工具时，都问到使用的方法，下面我来简单介绍 Google 网站管理员工具的基本功能。

目标地理区域：用于设置网站所面向的用户，国别级域名将自动关联为相关国家，国际域名可以手工设置目标地理区域。

首选域：如果你的域名同时使用了带 www 的和不带 www 的，就需要用首选域功能，设置一个在谷歌显示的域名，通常选择的是带 www 的域名，如果不设置，Google 会自动选择一个域名，但不利于进行搜索引擎的排名。

图片搜索：Google Image Labeler 是谷歌开发的一个可以帮助自己更好地识别图片中内容的一个程序，目前与搜索排名无关，不作详细讲解。

抓取速度：这个功能可以改变 Google 蜘蛛来访的频率，如果你的网站更新非常频繁，则

可以自定义抓取速度；如果你的网站只是一个小企业站，则建议让 Google 来确定抓取速度，否则太多的蜘蛛来到你的网站只会给你服务器增加负担，不会给你有什么帮助的。

网络抓取：在此处可以看到详细的 Google 蜘蛛在抓取网站内容时的错误反馈。

移动抓取：该功能作用不祥，暂时用不到。

内容分析：用于检测抓取网站时是否存在潜在的内容问题，包括重复内容、内容缺失或有问题的标题标记及元描述等。

热门搜索查询：可以观察到近期搜索进入你网站的关键词。因为该功能不是很准确，所以完全可以被专业的统计网站所取代。

Googlebot 看到的内容：这些统计信息向你显示了 Google 蜘蛛是如何查看你的网站的，分析了你的外链文字以及网站中的其他信息。

2. 统计信息介绍

抓取统计信息：以数据和波形图片的形式展示了 90 天内网站的抓取情况，以及 PR 的情况。

索引统计信息：介绍了几个相关的查询命令。

订阅用户统计信息：如果你的网站发布关于供稿的内容，此页会显示使用 iGoogle、Google 阅读器或 Orkut 等 Google 产品订阅这些供稿的用户数。

含外部链接的网页：展示所有拥有外部链接的页面，以及外部链接数量，点击“外部链接”列中的数字，可查看指向页面的链接示例列表。是分析外链的绝佳帮手，强烈推荐使用。

网站链接：这里展示的内容我们称为 Sitelinks，这可不是每个网站都有的，如果你拥有 Sitelinks，那就可以在这里调整显示的相关链接。

Sitempas：用来提交符合 Google 标准的网站地图，可以帮助 Google 抓取你的网站。

3. 工具的使用

分析 robots.txt：用来测试你的 robots.txt 文件设置是否正确，在上传前，可以先用该功能进行模拟测试，检查完毕后，会有反馈结果。

生成 robots.txt：如果你不熟悉 robots.txt 的规则，则可通过该功能生成 robots.txt 文件，操作比较简单。

网站管理员验证：该功能在开始使用 Google 网站管理员工具时已经验证了。

删除网址：如果谷歌收录了你网站中不希望收录的页面，可利用该功能快速提交删除申请，谷歌将在 3～5 天内删除索引。注意，该功能要慎用，如果操作不当，有可能整个网站的索引都被删除了。

增强型 404 页面：这里是 Google 给用户推荐的一个 404 页面，可以有效地提升用户体验，大家可以考虑使用。

小工具：可以将 Google 网站管理员工具集成到 Google 的主页，这个是 Google 的 iGoogle 功能，也就是自定义 Google 主页，可以方便站长观察自己网站的情况。

10.3 SiteMap 生成工具

网站地图（SiteMap）即网站管理者告知搜索引擎自身网站上有哪些可供抓取的网页。其中最简单的 SitepMap 形式，就是 XML 文件，在其中列出了网站中的网址及关于每个网址的其他元数据（上次更新的时间、更改的频率及相对于网站上其他网址的重要程度等），以便搜索引擎可以更加智能地抓取网站。

网络抓取工具通常会通过网站内部和其他网站上的链接查找网页。

SiteMap 会提供此数据以便允许支持 SiteMap 的抓取工具抓取 SiteMap 提供的所有网址，并了解使用相关元数据的网址。使用 SiteMap 协议并不能保证网页会包含在搜索引擎中，但可以向网络抓取工具提供一些提示以便它们更有效地抓取网站。

Sitemap 0.90 是依据创意公用授权，以相同方式共享（Attribution-ShareAlike Creative Commons License）的条款，被广泛采用，并得到 Google、Yahoo! 和 Microsoft 在内的众多厂商的支持。

Google SiteMap Protocol 是 Google 自己推出的一种站点地图协议，此协议文件基于早期的robots.txt文件协议，并有所升级。在 Google 官方指南中指出加入了 Google SiteMap 文件的网站将更有利于 Google 网页蜘蛛的爬行索引，这样将提高索引网站内容的效率和准确度。文件协议应用了简单的 XML 格式，一共用到 6 个标签，其中关键标签包括链接地址、更新时间、更新频率和索引优先权。

Google SiteMap 文件生成后格式如下：

```
<urlset xmlns= "网页列表地址" >
<url>
<loc>网址</loc>
<lastmod>2005-06-03T04:20-08:00</lastmod>
<changefreq>always</changefreq>
<priority>1.0</priority>
</url>
<url>
<loc>网址</loc>
<lastmod>2005-06-02T20:20:36Z</lastmod>
<changefreq>daily</changefreq>
<priority>0.8</priority>
</url>
</urlset>
```

百度 SiteMap 文件生成后格式如下：

```
<? xml version="1.0" encoding="UTF-8"? >
<urlset>
<url>
```

```
<loc>网页地址</loc>
<lastmod>2010-01-01</lastmod>
<changefreq>daily</changefreq>
<priority>1.0</priority>
</url>
</urlset>
XML 标签
changefreq:页面内容更新频率
lastmod:页面最后修改时间
loc:页面永久链接地址
priority:相对于其他页面的优先权
url:相对于前 4 个标签的父标签
urlset:相对于前 5 个标签的父标签
将一句一句讲解 xml 文件的每一个标签:
<urlset xmlns="
该行定义了此 xml 文件的命名空间，相当于网页文件中的<html>标签。<url>
```

这是具体某一个链接的定义入口，你所希望展示在 SiteMap 文件中的每一个链接都必须包含在<url>和</url>中。

```
<loc></loc>
```

用<loc>描述出具体的链接地址，这里需要注意的是，链接地址中的一些特殊字符必须转换为 XML（HTML）定义的转义字符，如表 10-1 所示。

表 10-1　转义字符的例子

字符	转义后的字符
符号&	&
单引号'	'
双引号"	"
大于号>	>
小于号<	<

```
<lastmod>2005-06-03T04:20:32-08:00</lastmod>
```

<lastmod>是用来指定该链接的最后更新时间，这个很重要。Google 的机器人会在索引此链接前先和上次索引记录的最后更新时间进行比较，如果时间一样就会跳过不再索引。所以如果你的链接内容基于上次 Google 索引时的内容有所改变，应该更新该时间，让 Google 下次索引时会重新对该链接内容进行分析和提取关键字。这里必须用 ISO 8601 中指定的时间格式进行描述，格式化的时间格式如下：

- 年：YYYY（2005）

- 年和月：YYYY-MM（2005-06）
- 年月日：YYYY-MM-DD（2005-06-04）
- 年月日小时分钟：YYYY-MM-DDThh:mmTZD（2005-06-04T10:37+08:00）
- 年月日小时分钟秒：YYYY-MM-DDThh:mmTZD（2005-06-04T10:37:30+08:00）

这里的 TZD 指定的是本地时间区域标记，如中国就是+08:00。

```
<changefreq>always</changefreq>
```

以上这个标签告诉 Google 此链接可能会出现的更新频率，如首页肯定就要用 always（经常），而对于很久前的链接或者不再更新内容的链接就可以用 yearly（每年）。这里可以用来描述的单词有 always、hourly、daily、weekly、monthly 和 yearly，其具体含义看单词的意思应该就明白了。

```
<priority>1.0</priority>
```

<priority>是用来指定此链接相对于其他链接的优先权比值，此值定于 0.0 ~ 1.0 之间。

此外，还有</url>和</urlset>，这两个是用来关闭 xml 标签的，其与 HTML 中的</body>和</html>是一个道理

另外需要注意的是，这个 xml 文件必须是 UTF-8 的编码格式，不管是手动生成还是通过代码生成，建议最好检查一下 xml 文件是否是 UTF-8 编码，最简单的方法就是用记事本打开 xml 然后另存为时选择编码（或转换器）为 UTF-8。

登录 Google，登录后转到 Your Sitemaps 状态页面，可以点击 Add a Sitemap + 跳转到提交页面进行 SiteMap 文件的提交。建议将文件放在站点根目录下。给 Google 提交你的 SiteMap URL 后可以看见在列表里已存在，不过这时候还没有生效，必须过几个小时后 Status 栏变成 OK，则表示正式生效，如果不是 OK，可以查看 Google 给出的状态标示查看是什么原因。

10.3.1　SiteMap 的在线生成步骤

百度对已提交的数据，不保证一定会抓取及索引所有网址，并且不保证其排名。

使用 Google SiteMap 能提高网站、网页在 SERP 中的排名（或提高 SEO 效果）SiteMaps 只是在一定程度上提高网站页面被收录、索引的效率，从这个意义上说，对网站的整体 SEO 效果是有所助益的。

但是 SiteMaps 与网站、网页最终出现在 SERP 中的排名则没有任何直接意义上的联系，这是两个虽有联系但实现过程中交叉部分又很少的两个机制。当然，如果非要强辩说，随着网站内页面收录数量增加，站内的交叉链接权值会相应提高，而这会对最终的排名产生影响，这也能在某种程度上说得过去，不过，影响多大就难说了。具体地图的生成步骤如下。

步骤01 打开浏览器输入在线生成 SiteMap 网站的网址。

步骤02 在表单中填入要生成 SiteMap 的网页的网址，然后点击提交。

步骤03 打开了生成的数据结果页面，复制文本框中的代码。

步骤04 新建一个文本文件把代码粘贴进去，然后另存为 UTF-8 格式的文件，文件名为 sitemap.xml，然后把这个文件上传到自己网站对应的根目录下。

步骤05 打开浏览器输入网址，点击右上角的登录，用自己的 Google 账户登录，还没有 Google 账户的用户先创建一个账户，然后再登录。

步骤06 登录成功后，点击进入用户管理中心，然后点击网站管理员工具。

步骤07 首先添加你的网址链接，添加成功后，在网址的对应项后面点击添加，打开 SiteMap 的添加页面，选择下拉菜单中的普通网站地图，下面出现文本表单，在空白表单后面添加 sitemap.xml 文件，然后点击提交。

步骤08 SiteMap 提交成功，等待 5 个小时之内 Google 收录你的网站，至于百度蜘蛛则是 html 格式的文件。

10.3.2 SiteMap 的使用

SiteMap 网站地图的使用目的就是为了方便搜索引擎蜘蛛抓取页面，如果 SiteMap 的连接中存在死链或坏链，则肯定会影响网站在搜索引擎中的权重，所以要仔细检查有无错误的链接地址，提交前通过站长工具，检查网站的链接是否有效。下面来介绍如何正确使用 SiteMap。

步骤01 添加网站，验证网站归属，等待审核。

登录后点击“添加网站”按钮。若网站删除或更换域名，则须重新提交。

点击“验证此网站”按钮，选择验证方式并获取验证文件，完成验证。

- 验证方式一：文件验证
- 验证方式二：HTML 标签验证

等待管理员审核，主要审核网站属性、网站内容质量等。审核时间最长可能需要一天。

步骤02 通过“数据管理”，添加新数据。

选择数据的类型，根据类型对应的 xml 格式规范部署你的数据文件，在地址栏中填写文件存放地址，点击提交即可，对文件处理时间的长短将视文件大小而定。

如果想在指定更新周期以外，手动通知我们数据有更新，可以在“数据管理”页面选中要手动更新的文件并点击“更新所选”按钮即可。

如果数据存在错误或不符合协议格式，则状态栏会显示为错误，请参考提示进行修改并

更新。

你可通过平台查看数据的统计信息，包括已抓取数量，最新处理时间等。

10.3.3 SiteMap 网站地图的注意事项

在 SiteMap 网站地图的使用中还有一些需要 SEO 人员注意的事项，如果处理不当也会带来不好的影响。下面给大家指出几点必须注意的事项。

1. 全部链接真实有效

地图的主要目的是方便搜索引擎蜘蛛抓取，如果地图存在死链或坏链，会影响网站在搜索引擎中的权重，所以要仔细检查有无错误的链接地址，提交前通过站长工具，检查网站的链接是否有效。

2. 简化网站地图

网站地图不要出现重复的链接，要采用标准 W3C 格式的地图文件，布局要简洁、清晰，如果是内容式地图，则每页不要超过 100 个内容链接，采用分页的形式逐一开来，这样方便搜索引擎蜘蛛逐页爬行。

3. 更新网站地图

建议经常更新网站地图，经常更新地图，便于培养搜索引擎蜘蛛爬行的粘度。经常有新的地图内容生成，长此以往，蜘蛛就会更加关注，也能更好地培养蜘蛛的爬行规则，这样网站内容能更快的被搜索引擎抓取收录。

10.4 Free Monitor for Google

Free Monitor for Google是一款免费的 Google关键词排名批量查询工具，和之前介绍的PR值批量查询工具 PaRaMeter都是 cleverstat.com 的产品。Free Monitor for Google 非常的小巧，只是查询排名的速度很慢。如果想不弹出验证码就能批量查询关键词排名，只能每次查询的时间间隔至少在 7 秒以上。

Free Monitor for Google 支持多国语言，但很遗憾，它不支持中文。不过其功能很简单不需要复杂的设置，稍微懂点英语的就能上手。

Free Monitor for Google 有以下特色。

- 可批量导入导出关键词。TXT 格式一行一个关键词。
- 支持多个网站，但一次只能批量查询一个网站。
- 可以设置代理，Google Web API。

- 可将排名结果以 CSV 格式导出。
- 支持多国语言，但无中文。

Free Monitor for Google 有以下缺点。

- 只能查询 Google.com 的排名。
- 只能用来查询英文关键词，中文关键词有点问题，多了就崩溃。

综上所述，若你是做 Google.com 的英文排名，又不想花钱买软件的话，Free Monitor for Google 则很适合，如图 10-7 所示。

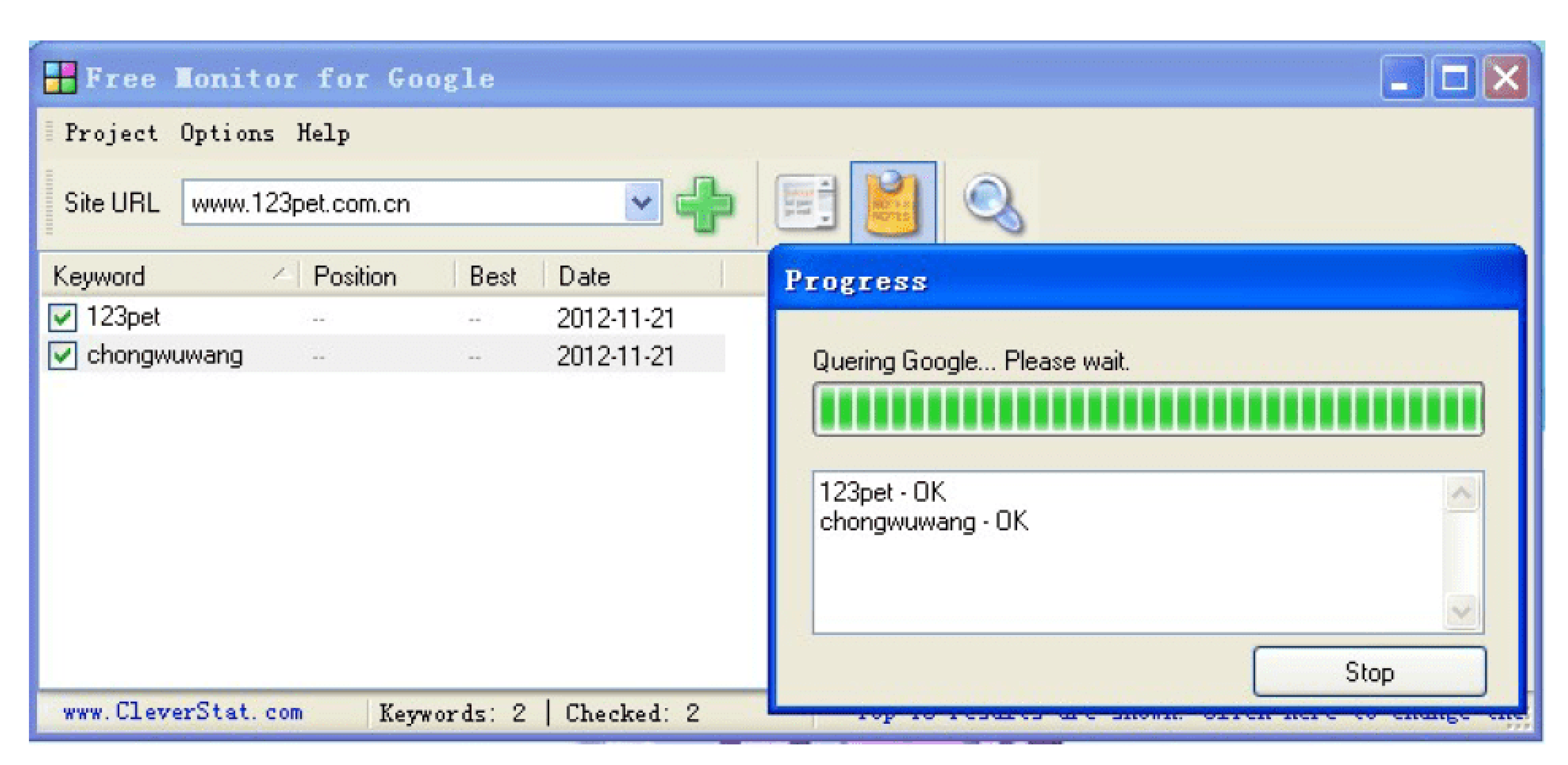


图 10-7　Free Monitor for Google 批量查询后的显示界面

10.5 整合搜索优化

整合搜索于 2007 年年底由谷歌首先推出，随后被所有主流搜索引擎所采用。所谓整合搜索，就是在正常搜索结果页面中同时显示了文字页面之外的图片、视频、新闻、博客、地图、图书等垂直搜索结果。其实所有主流搜索引擎早就推出了整合搜索，用户只要单击搜索结果页面上方的垂直搜索导航，就可以显示出相应的结果，如图 10-8 所示。



图 10-8　谷歌整合搜索

不过人都是懒惰的，搜索引擎发现绝大部分用户很少点击垂直搜索导航，所以就把垂直搜索的结果整合到正常网页搜索结果页面中了，这样对 SEOer 来说既是一个机会，也是一种挑战。

说它是机会的原因在于，一些中小型网站在文字页面没机会排到前两页时，却有机会通过整合搜索结果进入到前几名。说它是挑战是因为，传统 SEO 要优化的只是普通页面，以文字为主。而现在 SEO 的工作范围更加广泛，各种多媒体内容和社会化媒体内容都要被纳入到 SEO 的范围。SEO 人员必须从战略思考上提高一个层次，不再局限于传统意义上的页面优化。

利用新闻搜索进行优化，加入搜索引擎新闻搜索栏最主要的一点是要想办法被搜索引擎纳入新闻源。搜索引擎返回资讯内容时，先选定哪几条新闻应该被显示，这主要是靠用户关注度以及媒体报道数量判断的。如果想让自己的新闻在结果中排名靠前，需要注意以下几点。

1. 发布和更新时间

由于新闻内容本身的特性，越新发布的内容越有排名优势。这就要求你要有敏锐的嗅觉，能够在第一时间察觉到新闻的发生，并把它写成文章发布到网上。

2. 原创

转载其他新闻内容，效果没有原创的那么好。普通网页都要求原创，这种非普通网页的结果就更不用说了。除非你的网站自身权重比较高，这样还可以获得一个比较不错的排名。

3. 地域相关

很多新闻有地域性，同一个地区的新闻源网站具有优势（这个与用户在不同的地方搜索同一个词显示的结果却不同是一个道理）。如这则新闻是发生在武汉的，那么武汉的新闻网在这则新闻的排名上肯定具有地域优势。

4. 标明时间、地点

新闻稿中最前面标明时间、地点是标准的新闻写作格式。以这样的格式呈现出来的新闻，也有助于搜索引擎判断新闻内容的地域性和时效性。

5. 网站权重

这里所说的权重与 PR 值或其他来源外部链接的权重不同，更多的是与主题相关的权重。如这则新闻是关于体育的，而你的网站上大部分新闻内容都是与体育相关的，这样你的网站就会有一个不错的权重。

6. 点击率

同样的新闻内容，点击率高的那一条排名会飞快地上升。这就需要你想办法吸引更多的人来观看你的新闻，并把这则新闻分享出去，让更多的人点击和观看。

7. 图片

如果在你的新闻页面中配上图片，则有助于吸引目光，提高点击率。

8. 其他页面因素

与其他普通页面相似，页面的标题、URL 及正文中出现的目标关键词等，都会计算到排名中，所以这些要素也不能轻视。

10.6 小结

能够合理地应用工具可以让站长们用更短的时间做比对手更多的事情，也会让我们的工作达到事半功倍的效果，通过 Google 管理员工具、百度指数、站长工具等一些 SEO 辅助工具可以使我们轻易地掌握自己网站的收录及反向链接等重要信息，为网站优化提供了数据支持。

第 11 章　经典案例分析

分析别人的网站也是学习 SEO 的技巧之一，通过分析竞争对手的网站来选定自己网站的优化方向是很多站长采用的方法，也是相当行之有效的。对排名较好的网站进行细致地 SEO 分析绝对会给你带来意外的收获。本章选择了一些比较有代表性的网站来分析其优化的利与弊。

11.1　123 宠物网

123 宠物网，成立于 2009 年 2 月 5 日。以丰富的专业知识、轻松愉悦的氛围和独具特色的内容吸引了大量爱宠物人士在此驻足停留、安家落户，该网站正以惊人的速度发展壮大，成为中国目前最具专业的狗主人社区及宠物狗交易市场。

1. 网站基本信息

网站名称：123 宠物网

百度收录：11 万以上

反链：19 万 8 千

网址：www.123pet.com.cn

二级域名：bbs.123pet.com.cn、product.123pet.com.cn/、blog.123pet.com.cn

流量：123 宠物网的平均日浏览量约为 9 万左右，信息版块和论坛版块通过百度关键词带来的 IP 每日约有 2 万左右，如图 11-1 所示。

☑ PV ☐ 独立访客 ☐ IP ☐ 新独立访客

日访问量分布　　导出此表数据

日期	PV	独立访客	IP	新独立访客	人均浏览次数	比例
总计	2800579	2142688	834496	1111690	1.31	-
2011-09-28	94069	72898	27478	13350	1.29	3.36%
2011-09-27	94930	76663	28188	13390	1.24	3.39%
2011-09-26	104592	86258	30201	41827	1.21	3.73%
2011-09-25星期天	100348	83698	29569	48097	1.20	3.58%
2011-09-24星期六	98601	83462	29150	48763	1.18	3.52%
2011-09-23	95763	78211	28001	37148	1.22	3.42%
2011-09-22	92663	73266	27188	39698	1.26	3.31%
2011-09-21	94440	75463	27865	45011	1.25	3.37%
2011-09-20	106205	85853	29777	46070	1.24	3.79%
2011-09-19	110361	90323	30548	51061	1.22	3.94%
2011-09-18星期天	109813	89956	30809	41891	1.22	3.92%
2011-09-17星期六	104096	87729	30309	53318	1.19	3.72%

图 11-1　123 宠物网 SEO 流量分析

百度指数主域名相关关键词，如图 11-2 所示。


百度权重查询

HTTP:// www.123pet.com.cn 默认排序 查询

百度权重为 5，共找到 1060 条记录，预计从百度来访 5574 次 ，子域名：当前:www.123pet.com.cn

序号	关键字	指数	排名	网页标题
1	马犬价格	167	1	马犬价格|马犬多少钱|马犬图片-狗市-123宠物网
2	狐狸狗价格	112	1	狐狸狗价格|纯种狐狸狗多少钱|狐狸狗图片|狐狸狗买卖-狗市-123宠...
3	小狗名字	103	1	小狗名字大全(中文狗名大全,英文狗名大全) - 123宠物网
4	腊肠狗价格	96	1	腊肠狗价格|纯种腊肠狗多少钱|腊肠狗图片|腊肠狗买卖-狗市-123宠...
5	巴西龟的寿命	68	1	巴西龟的寿命有多长
6	兰花多少钱	67	1	兰花|兰花价格|兰花多少钱|兰花图片 - 123宠物网
7	八哥犬多少钱一只	59	1	巴哥犬巴哥犬价格八哥犬多少钱一只巴哥犬图片八哥犬123宠
8	小狗的名字	402	2	小狗名字大全(中文狗名大全,英文狗名大全) - 123宠物网
9	马犬图片	293	2	马犬价格|马犬多少钱|马犬图片-狗市-123宠物网
10	杜高犬价格	165	2	杜高犬价格|杜高犬多少钱|杜高犬图片-狗市-123宠物网
11	狐狸狗图片	142	2	狐狸狗价格|纯种狐狸狗多少钱|狐狸狗图片|狐狸狗买卖-狗市-123宠...
12	买猫	121	2	北京朝阳买猫|卖猫网朝阳卖猫的地方 - 123宠物网
13	狗的生活习性	101	2	小狗的生活习性
14	宠物兔价格	98	2	宠物兔|宠物兔价格|宠物兔多少钱|宠物兔图片 - 123宠物网

图 11-2　123 宠物网主域名关键词查询

域名注册日期：2009 年 05 月 15 日

推广：百度知道、SEO 资料站等，有着大量的反链，且其内容更新量大、原创较多。

2. 网站 SEO 分析

从理想状态上讲，应该是站点性质决定关键词策略，然后关键词策略决定网站结构及网站内容。但实际情况通常是，一个网站的结构及网站的内容是有固定套路的，如 B2C 商城网站就是有固定的结构，B2C 商城基本上就是首页、类目页及单品页，如果你改变了这种结构，就会给用户造成浏览的障碍，用户学习成本较高。因此，实际操作中，通常是由现有的网站结构来反推关键词策略。

```
<title>123 宠物网—中国最专业的狗主人社区及宠物狗交易市场</title>

<meta name="keywords" content="宠物狗社区,宠物狗交易,宠物狗狗,宠物狗,123 宠物网,狗狗图片,宠物网" />

<meta name="description" content="123 宠物网是中国宠物行业最具专业的宠物狗主人社区平台,中国最具专业的宠物狗交易市场,在这里你可以交流养宠物狗狗的心得,在论坛、家园结交狗友并分享学习养宠物狗狗的知识,享受最具乐趣的线上线下活动,同时可与专业的犬舍进行宠物狗买卖,宠物狗交易" />
```

关键词分布。网站右上角放置主要犬种文字导航链接，如图 11-3 所示。



图 11-3 123 宠物网犬种导航链接

网站底部放置犬种辅助导航链接，如图 11-4 所示。



图 11-4 网站底部犬种辅助导航链接

论坛版块首页二级辅助导航，如图 11-5 所示。



图 11-5 论坛版块二级辅助导航

点击左上角的“123 宠物网为你提供最全犬种，宠友们记得收藏！”，其实际上是“收藏本站”，便于客户把网站收藏起来，下次再看。

网站底部的“联系我们”，有着各部门的电话、传真、手机、QQ 和邮箱，把这些联系方式统一放在一个页面。

顶部的导航，包括宠物狗、哈士奇、藏獒、比熊犬、萨摩耶、金毛犬、高加索犬、吉娃娃、松狮犬、贵宾犬等，主要犬种关键词按照用户浏览习惯放置在网站右上角的位置，同时也可增加这些页面的权重。

使用最新资讯的滚动更新文本框来使得首页页面常有更新，如图 11-6 所示。

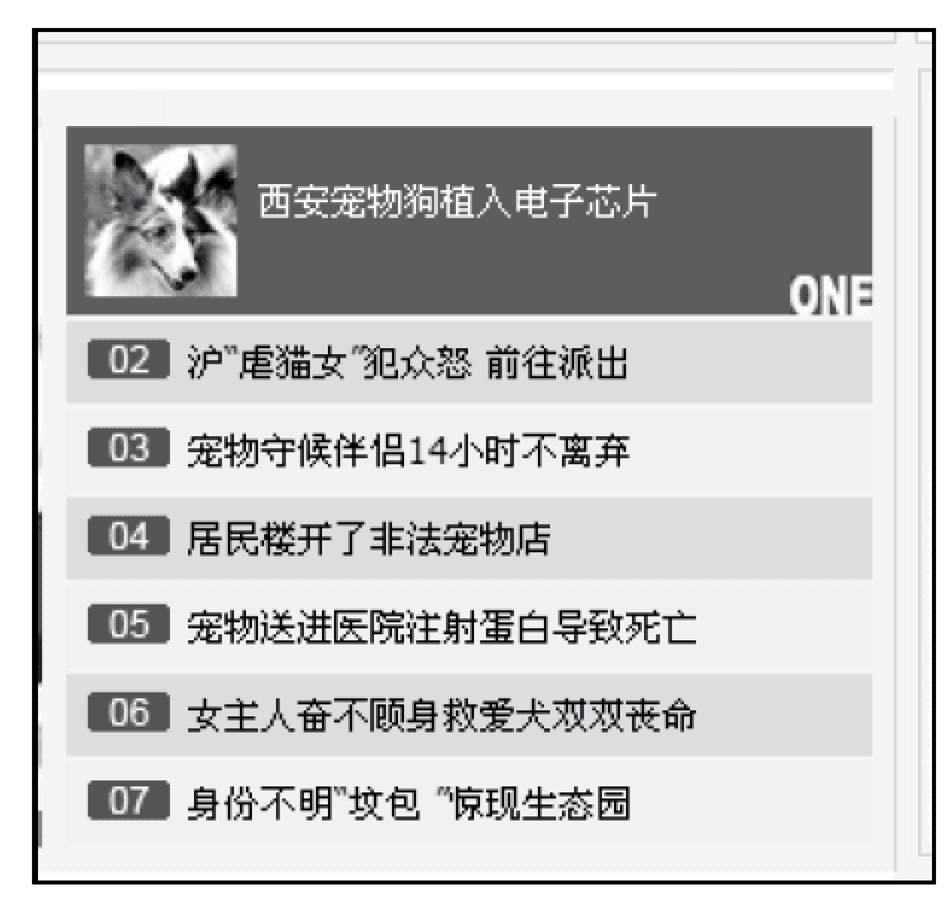


图 11-6　首页滚动资讯文本框

首页底部，不定期的交换高质量并且均为宠物行业的友情连接。再加上网站备案信息、公司名称、联系方式等信息，在将访客可信度提升的同时，搜索引擎也很关注这些细节。

11.2 女性网站的优化分析

一般来说，网站首页基本上是主攻品牌词，单品页基本上就是产品名称的关键词，如果是某些标准化的产品，如图书、3C，那么单品页还会竞争一些诸如 ISBN、产品型号等长尾关键词。而类目页的关键词策略则根据网站产品的不同而不同，基本上有两种做法：一种是如果该网站售卖的是别人品牌的产品，则关键词策略基本上是“品牌+产品类目”的排列组合；另外一种就是网站售卖的都是自己的产品，因为都是自家产品，品牌关键字都是自家的，这种排列组合的形式就不适用。梦芭莎卖的产品也都是自己的品牌，那么，它是如何做类目页的关键词策略的呢？如图 11-7 所示。

历史	Url	Keywords	百度排名网址	百度排名	谷歌排名网址	谷歌排名
1	www.moonbasa.com	挺拔胸罩	www.moonbasa.com/Pro...	1/2	www.moonbasa.com/pro...	4
2	www.moonbasa.com	集中款内衣	www.moonbasa.com/Pro...	1	www.moonbasa.com/pro...	1
3	www.moonbasa.com	魔术内衣	www.moonbasa.com/Pro...	3/1	www.moonbasa.com/pro...	25
4	www.moonbasa.com	轻薄内衣	www.moonbasa.com/Pro...	3/2	www.moonbasa.com/pro...	14
5	www.moonbasa.com	透气文胸	www.moonbasa.com/Pro...	2/1	www.moonbasa.com/pro...	1
6	www.moonbasa.com	小可爱文胸	www.moonbasa.com/Pro...	3/1	www.moonbasa.com/pro...	1
7	www.moonbasa.com	a 罩杯有多大	www.moonbasa.com/Pro...	3	www.moonbasa.com/pro...	8
8	www.moonbasa.com	b 罩杯有多大	www.moonbasa.com/Pro...	5	www.moonbasa.com/pro...	8
9	www.moonbasa.com	双排扣内衣	www.moonbasa.com/Pro...	3/9	www.moonbasa.com/pro...	1
10	www.moonbasa.com	四排扣内衣	www.moonbasa.com/Pro...	3/1	www.moonbasa.com/pro...	1
11	www.moonbasa.com	透明肩带	www.moonbasa.com/Pro...	3/89	www.moonbasa.com/pro...	25

图 11-7　关键词地址目录

上图中，选择了梦芭莎的“文胸”类目作为例子，文胸这个类目的筛选条件为“按系列、材质、按尺码、按罩杯、按排扣、按配饰”等，基本上这个类目的关键字策略就是以这些筛选条件为组合关键词。“品牌+产品类目”这样的关键字组合可以靠程序自动生成的，但是这种个性化的关键字设置则需要人工撰写，从梦芭莎的类目 title 的撰写上来看，梦芭莎的 SEO 人员在这方面是下了很多功夫的，每个类目基本上都是选取了检索量较高的关键字形式来进行设置。

另外，梦芭莎的 SEO 策略还有一个重要的组成部分，就是对于“产品+好不好”，“产品+怎么样”这样的词缀做了关键词策略规划，这种类型的关键字搜索量大、用户意向比较明确，而且由于是长尾关键词，竞争通常也不激烈，这种类型的关键词是长尾策略的最佳载体，如图 11-8 所示。

关键词	12.22	12.21	12.20	12.19	12.18	12.17	12.16	总计
怎么样	267	373	302	320	274	459	406	2401
让子弹飞怎么样	333	421	469	369	426	565	391	2974
非诚勿扰2怎么样	423	44	27	18	15	14	10	551
诺基亚5230怎么样	942	851	988	971	890	809	793	6244
诺基亚c6怎么样	913	1015	904	890	958	841	735	6256
诺基亚c7怎么样	1031	987	994	873	795	841	749	6270
申通快递怎么样	199	193	174	125	123	117	139	1070
魅族m9怎么样	841	709	784	637	461	505	475	4412
htc手机怎么样	877	833	772	675	720	775	752	5404
诺基亚x6怎么样	692	762	728	692	708	620	523	4725
诺基亚n8怎么样	555	598	558	495	464	467	500	3637
华硕笔记本怎么样	235	228	289	269	215	217	193	1646
相宜本草怎么样	965	978	928	767	809	837	730	6014
诺基亚c5 03怎么样	500	547	546	430	411	420	419	3273
联想笔记本怎么样	339	381	370	319	303	305	295	2312

图 11-8　关键字的搜索量

通常，这样的关键词排列组合都是站长们（导购网站表现得特别明显）利用 SEO 截留官网流量的典型做法，不过梦芭莎这次好像没有给这些站长机会，如图 11-9 所示。

历史	Url	Keywords	百度排名网址	百度排名	谷歌排名网址	谷歌排名
1	review.moonbasa.com	波普几何女鞋	review.moonbasa.com/...	1	--	--
2	review.moonbasa.com	波普几何女鞋怎么样	review.moonbasa.com/...	1	--	--
3	review.moonbasa.com	波普几何女鞋好不好	review.moonbasa.com/...	1	review.moonbasa.com/...	15
4	review.moonbasa.com	镂空毛织披肩	review.moonbasa.com/...	1/1	--	--
5	review.moonbasa.com	镂空毛织披肩怎么样	review.moonbasa.com/...	1	--	--
6	review.moonbasa.com	镂空毛织披肩好不好	review.moonbasa.com/...	1	--	--
7	review.moonbasa.com	黑白方格毛围巾	review.moonbasa.com/...	1	--	--
8	review.moonbasa.com	黑白方格毛围巾怎么样	review.moonbasa.com/...	1	--	--
9	review.moonbasa.com	黑白方格毛围巾好不好	review.moonbasa.com/...	1	--	--
10	review.moonbasa.com	玫瑰情人围巾	review.moonbasa.com/...	3	review.moonbasa.com/...	17
11	review.moonbasa.com	玫瑰情人围巾怎么样	review.moonbasa.com/...	2	--	--
12	review.moonbasa.com	玫瑰情人围巾好不好	review.moonbasa.com/...	2	--	--
13	review.moonbasa.com	光面提花素雅文胸	review.moonbasa.com/...	1/1	--	--
14	review.moonbasa.com	光面提花素雅文胸怎么样	review.moonbasa.com/...	1	review.moonbasa.com/...	1
15	review.moonbasa.com	光面提花素雅文胸好不好	review.moonbasa.com/...	1	review.moonbasa.com/...	1

图 11-9　梦芭莎关键词排名

虽然这些长尾的关键词检索量不高，但胜在数量多，而且这些关键词的意向十分明确，对于网站吸引长尾关键词流量，效果非常明显。

另外，梦芭莎似乎也想抢夺竞争对手的关键词，但似乎效果还没完全展示出来。

11.2.1 页面优化

确立了关键词策略，下一步就要考虑如何将这些策略贯彻到网站结构及页面要素上。梦芭莎在这方面的操作倒是中规中矩，基本上就是个性化地设置页面的 meta 元素。另外，为了凸显页面的核心关键字，在页面的边栏上增加了一个文本块，如图 11-10 所示。

B罩杯文胸B罩杯有多大

梦芭莎B罩杯内衣，自然上托推高，有效锁定脂肪，具有侧推聚拢，使胸部更集中的功效，使胸型更完美，立显丰盈、圆润女性曲线之美

图 11-10　文本块

如在“梦芭莎文胸-按尺码大小查看-B 罩杯文胸”类目页面中，梦芭莎的 SEO 人员就要在该页面边栏增加这样一个文本块，跟其他网站文本块不同的是，为了让这些文字更具有可读性，这些文本不是由程序生成的，而应该是由编辑人员手工撰写的，这个工作量很巨大，而且还要挖空心思把每个版块的内容都写得不一致，很是不容易。

另外，为了展示“产品+好不好”、“产品+怎么样”这样的关键词策略，梦芭莎开设了一个评论的子域 review.moonbasa.com，这个子域的开设十分巧妙，一方面，它将用户对商品的评论做了一个聚合，很好地体现了既定的关键词策略。另一方面，它也让 B2C 网站页面无法经常更新的难题有了一个较好的解决方案，借助这个子频道，很好地解决了 SEO 与用户体验的难题，同时也很好地利用了 UGC 的内容，这很值得电商们学习和借鉴。

从该频道在搜索引擎上收录的表现来看，搜索引擎也是十分喜欢这样的用户评价聚合页面，所以，商朋友们，如果你现在还在为网站收录情况不佳而苦恼的话，就赶紧做个用户评价聚合页面吧，加索引页是增加网站收录的不二选择。

11.2.2 外链策略

因为不是内部人员，所以无法跟大家分享他们的具体外链策略，只能借助工具查询梦芭莎的外链组成，猜测一下梦芭莎的外链策略。

（1）友情链接交换（这个不提）。

（2）付费链接

基本上现在的网站 SEO 或多或少都有使用付费链接这种手段，毕竟花费少、见效快，搜索引擎也比较难以分辨，而且性价比较高。试着分析了一下他们比较好的外链资源，不知道他们是以何种方式获得以下这些外链的。

- beauty.daqi.com
- g.pptv.com
- lady.163.com
- lady.qq.com
- women.daqi.com

（3）导航站

通过分析梦芭莎的外链构成，发现他们拥有大量导航站的链接，这些链接不是通常的直链，而是采用了：http://www.moonbasa.com/#a_id=1188-mzny 这种类型的链接，应该是他们网站购买了大量的导航站链接，后面的参数是跟踪订单来源的，目前搜索引擎对于 cps 链接对网站的加权作用并没有说得很明确，不过从实践上来看，这种外链的效果也不错。梦芭莎的导航站链接很多，这里仅摘取以下几个。

- www.hao123.com
- hao.360.cn
- www.114la.com
- www.sogou.com
- www.2345.com
- www.2345wz.com
- www.7255.com
- www.7322.com
- www.7345.com

（4）网站联盟

由于梦芭莎拥有自己的网站联盟，这也使得中小网站主、网络推手们有足够的动力去帮梦芭莎进行推广，因此在其外链构成中，也有不少的个人网站、个人博客、推广软文、小导航站放上了梦芭莎的网站推广链接，这也在客观上增加了梦芭莎的外链数量。

总的来说，梦芭莎在 SEO 方面做了不少细致的工作，其效果也十分明显，值得电商朋友们学习、借鉴。我也希望通过这个粗浅的剖析，能起到抛砖引玉的作用，让大家一起来探讨电商网站的 SEO 策略及方法。

11.3 天津租车网

天津租车网通过对百度关键词和 Google 关键词的分析之后，为了避免更激烈的竞争，也为了更快地产生效果，选择了定位更准的“天津租车”为核心词，“天津租车信息”、“天津租车服务”、“天津商务租车”、“天津租车网（站名）”为辅助词，其原因如下。

（1）公司租车业务基本集中再天津地区。

（2）从“租车”的百度相关关键词查询中能看到，大部分的补充词都是以地方分类的。

（3）“租车”这个关键词竞争很激烈，要在短期内获得高排名十分不容易。并且在百度的第一页基本都被推广“占据”了。

网站首页 title 将以上几个关键词列出：“天津租车、天津租车信息、天津租车服务、天津商务租车 - 天津租车网”。

当用户查到租车网之后，最关心的应该是租车价格、有没有想要的车型、公司信誉如何这几点。

所以再首页上，直接将这几项用户最关心的问题明显出来，问题题目用<h1>包括，题目中带关键词，把握关键词密度，不要过高，以免被认为是作弊。

页面代码书写。页面的代码书写是极其简单的，要通过 W3C 的任何标准认证都不是难事，即便有错误，稍微修改一下即可。直到现在网站也未进行过任何美化工作。

外部链接。非内容性的商业站要找高质量外部链接实在不容易，很多个人站都不愿意和笔者做链接。所以先在权重较高的站给租车网做了几个外部链接，然后到免费信息站发布了几篇软文和广告，文章中带有链接。

11.4 芒果网优化分析

从整个网站和介绍页面上看，芒果网主要是做商旅服务，有四个主打产品：酒店预订、机票预订、旅游套票和旅行管理。其主推的关键字：酒店预订、机票预订、旅游度假。

从搜索引擎各种数据、实际排名、网站结构、页面结构等方面进行分析，芒果网对搜索引擎的友好度简直可以用糟糕透顶来形容。这里主要分析了在百度的情况，其他搜索引擎也大概看了一下，没有太大差异，如 site 是 38600，domain 是 818，link 是 12。下面将从几个方面来分析芒果网的现状。

1. 产业竞争压力大

商旅这块是最近几年的热点，目前在各大搜索引擎中都有若干个热门关键词被竞价，百度最甚、雅虎其次、谷歌也有赞助商的连接。在这三个搜索引擎结果页里除去付费的以外，

前几名基本都让 e 龙和携程占据，前几页被各大旅游信息网站占据。就商旅行业来说，e 龙和携程已经占据相当大的市场分额，而且通过几年的积累，e 龙更是占领了大部分门户网站的旅游频道机票、酒店预定的位置。

2. 竞争对手做得更好

粗略的看了看 e 龙和携程的网站，无论从网站结构、页面结构、可用性、易用性、内容等方面都感觉相对比较完善。

3. 网站的构成

全站使用统一域名：mangocity.com；Web server 使用 Apache/2.0.52 （Red Hat）；页面使用.htm 和.jsp 两种文件。

4. 网站的文件结构混乱

整个网站的多层文件结构非常混乱，没有统一的规划。

- 积分商城：http://www.mangocity.com/Gift/index.jsp
- 机票：http://www.mangocity.com/TicketWeb/default.jsp

存在不少 URL 携带参数的动态页面。如：http://www.mangocity.com/MangoWeb/MainServlet？action=Web_HotelSearchPre。

通过 cms 发布的但最终页面被放置到了第五层内。如：http://www.mangocity.com/cms/2006-09/04/content_880.htm。

论坛的读版页及读贴页均为动态页面。

- http://www.mangocity.com/mangobbs/listTopic.html？boaId=3
- http://www.mangocity.com/mangobbs/showTopic.html？topicId=8618&boaId=3

存在跳转的页面。

- http://www.mangocity.com/package/package-list.htm
- http://www.mangocity.com/MangoPkgWeb/querySpecialSaleAction.do？operate=querySpecialSaleProduct&departCtiyId=70002&specialType=1

5. 网站的页面结构混乱

页面不符合 Web 标准，页面内 table 和 div 各类元素混杂，大量使用 table 元素进行排版，各类页面的信噪比都很低。

首页代码 1600 多行，实际的内容却非常有限，页面被大量的 JS、Flash、Map、Form 占据，其他如机票、酒店预定页都存在这个问题，非常不利于蜘蛛抓取。

- 页面 meta 混乱。

- 有的页面 title 顺序混乱甚至没有。
- 多数页面没有 Keywords 和 Description，有的页面都是一样的内容，而也有的页面内容的关联性并不大。
- 论坛页面存在两个 title。
- foot 内的链接竟然使用表单按钮。

6. 缺乏内容运营

没有专题，内容组织混乱，可以说是基本无序、没有相关推荐甚至缺乏站内连接。

11.5 深圳信息网

深圳，是一个对网络依赖度非常高的城市，造就了其本地网上分类信息的火暴。在激烈的竞争中，有一家网络公司引起了深圳人的注意，那就是深圳 160 信息网（后文简称深圳 160），其公司名为深圳汇深网信息科技有限公司。

11.5.1 背景分析

深圳 160 大概开始于 2003 年，从 2004 年底，笔者开始注意到了这家公司的操作模式，通过建立大量网站，以获取较好的搜索排名，并加上一定量的广告投入，来推广自己网站的本地信息，然后再拉拢商家在其中刊登分类广告。

对于深圳 160 的这种商业模式，主要是依赖于网站本身在搜索引擎的表现，然后再辅助以广告。所以如何获得大量的免费的自然搜索排名成了深圳 160 成功与失败的关键。

与此同时中国的 SEO 也开始在 2003 年后逐步兴起，在 2004 年底的时候，垃圾 SEO 已经盛行达到了一个顶点。深圳 160 也顺应时代潮流，投入了一大批网站，照着别人的优化手段开始了垃圾 SEO 之旅。优化的手段包括在网站制造大量的垃圾页面，使用跳转到处留言添加连接等手段。

然而从 2005 年开始 Google、百度等对 SEO 作弊进行了强烈的惩罚，深圳 160 也不列外。其主打网站 www.sz160.com 也就是在这一时期被 Google 删除的，同一时间还有其他多个网站被清除，如 ww.sun.org.cn。

深圳 160 网站也正是从这时开始反思，抛弃了中国 SEO 作弊常见的“经典手法”，转而开始了一场终极 SEO 作弊大策动。

曾经认为而且现在也认为深圳 160 的商业模式和理念是不错的，但是当笔者仔细分析其操作方法的时候，只能说一切太疯狂了。是笔者见过的投资成本最大，历时最久，动用人员最多的 SEO 作弊行为了。

而其作弊的方式非常隐蔽，但其手段却非常简单。就是制造大量的垃圾网站，不断地往

该网站堆积大量高度相关的内容。然后通过复制，再稍微调整结构、标题，改变域名，调换存放路径等，即可将一个网站变成上百个网站。他们不需要像其他 SEO 作弊那样寻找连接，因为他们本身就有足够的资源。虽然他们大量的网站类似，但一般人也很难发现，因为域名不同，网站名不同，很难让人鉴别出来。

换句话说，用深圳 160 的手法可以彻底的玩弄搜索引擎。如要优化每个行业的某些关键字，那么就先制作一个网站，管它行业里的什么内容，都往里面加。让这个网站的页面数量变得非常大，以确保有足够的页面内部连接。然后把这些内容的显示方式和网站结构稍微调整，让一个网站变成 10 个网站，让 1000 个页面迅速变成 10000 个页面，这时候搜索引擎不管怎么搜，按概率论算，轮也会轮到其中的一些页面排在搜索结果前面，加上针对的关键字几乎都是本地信息，竞争程度不一定非常激烈。所以很容易就可以在搜索引擎中占据大量的搜索排名。

11.5.2　操作手法抓图分析

接下来用抓图的手法来看一看他们的操作方法。

不同网站之间互为镜像页面新闻主题内容相同，结构相同。只是文件在网站中的位置不同、标题不同而已，如图 11-11 所示。

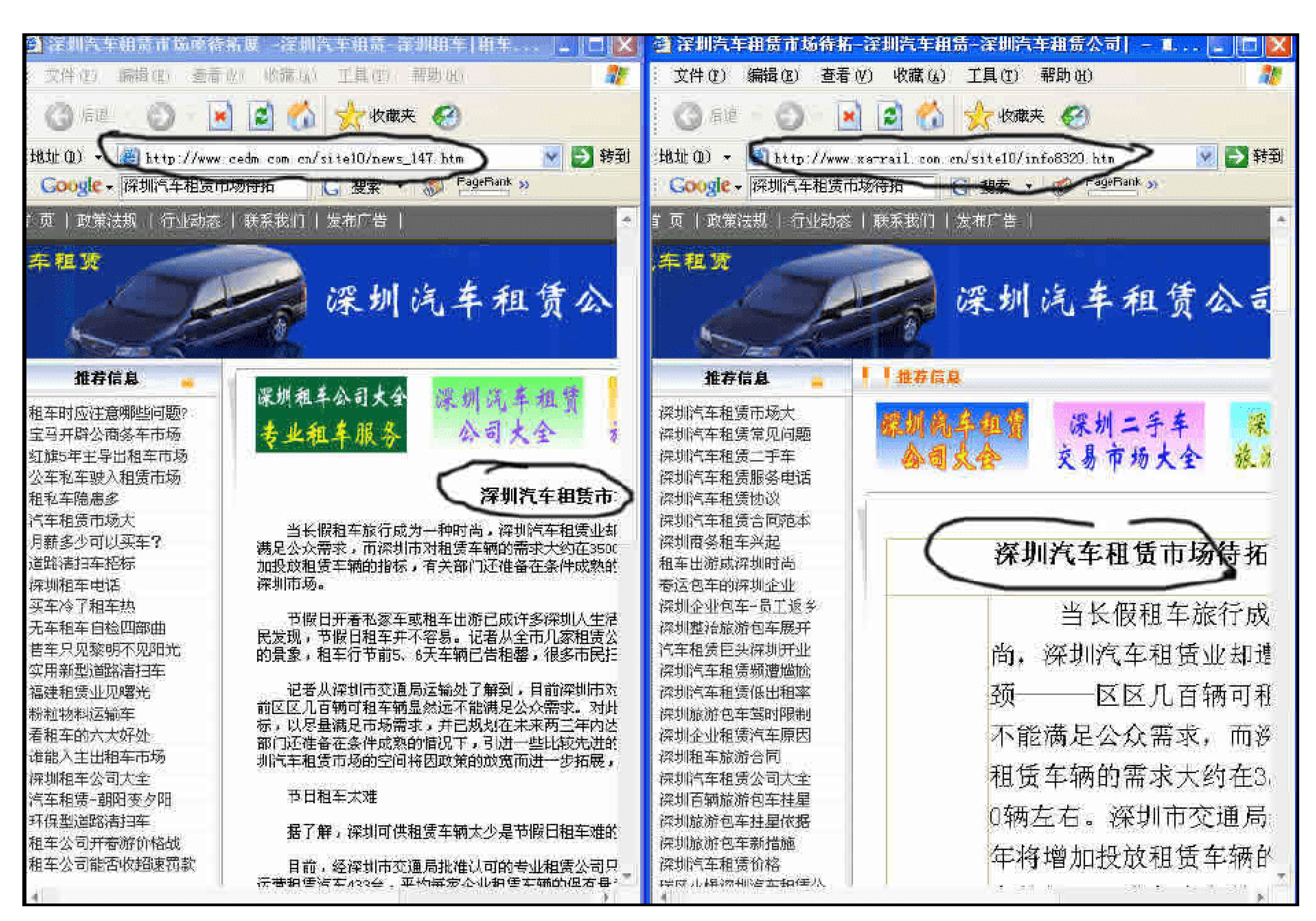


图 11-11　不同网站之间互为镜像

图 11-12 为同一网站中创建同样内容的镜像页面，只是对新闻标题取名不同。

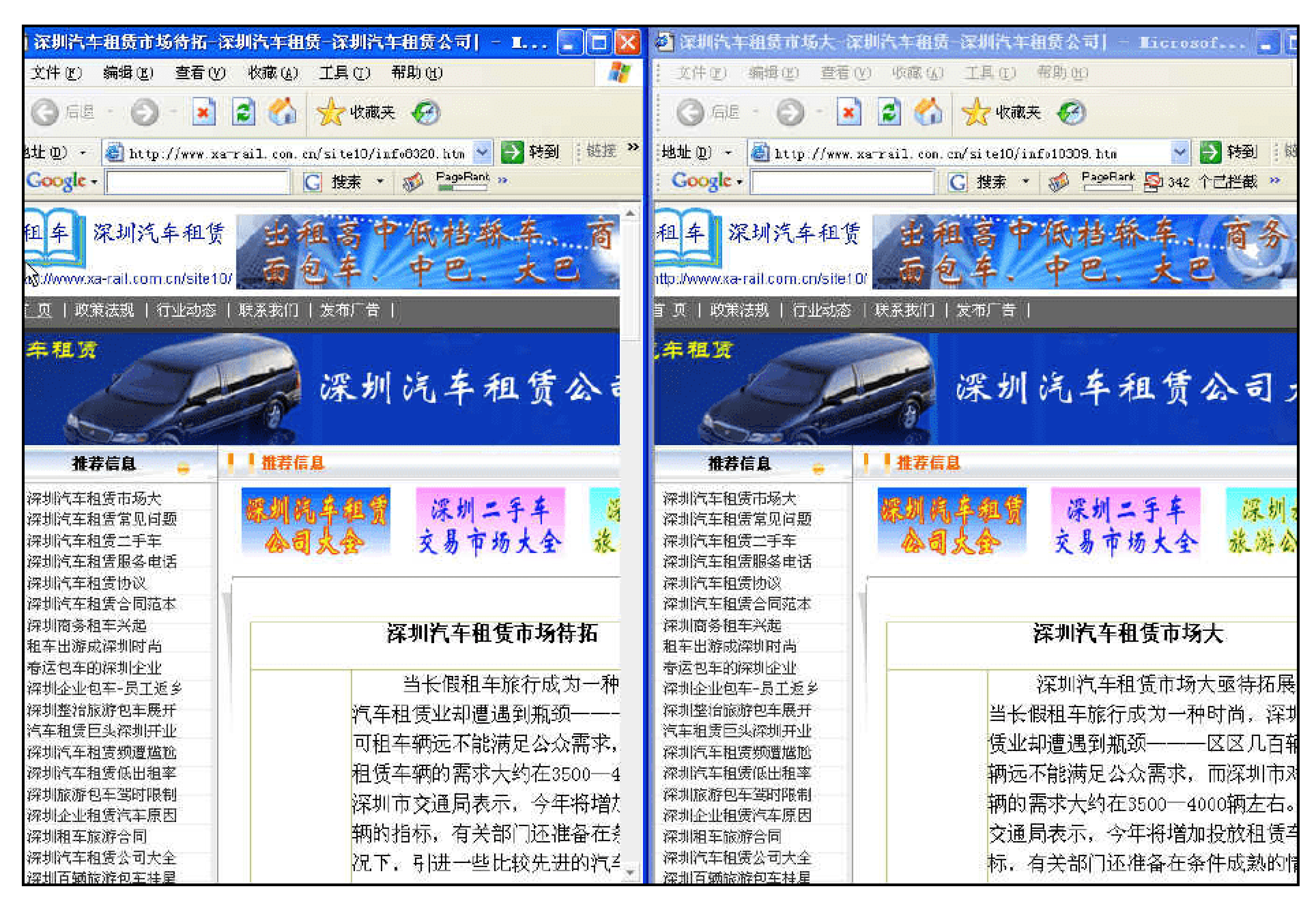


图 11-12　新闻标题取名不同

图 11-13 为两个网站的内容，结构完全一样。只是标题不一样，网站名称不一样，域名不一样。即将 www.smers.com.cn/site48/ 路径下的网站重新上传到另一域名 www.sinovel.com.cn 中。



图 11-13　网站名称不一样

图 11-14 为同一跳蚤市场，结构稍微调整，不同域名共用同一数据，网站命名不同。即将 www.smers.com.cn/site48/ 路径下的网站重新上传到另一域名 www.sinovel.com.cn，然后将网站改个名字。

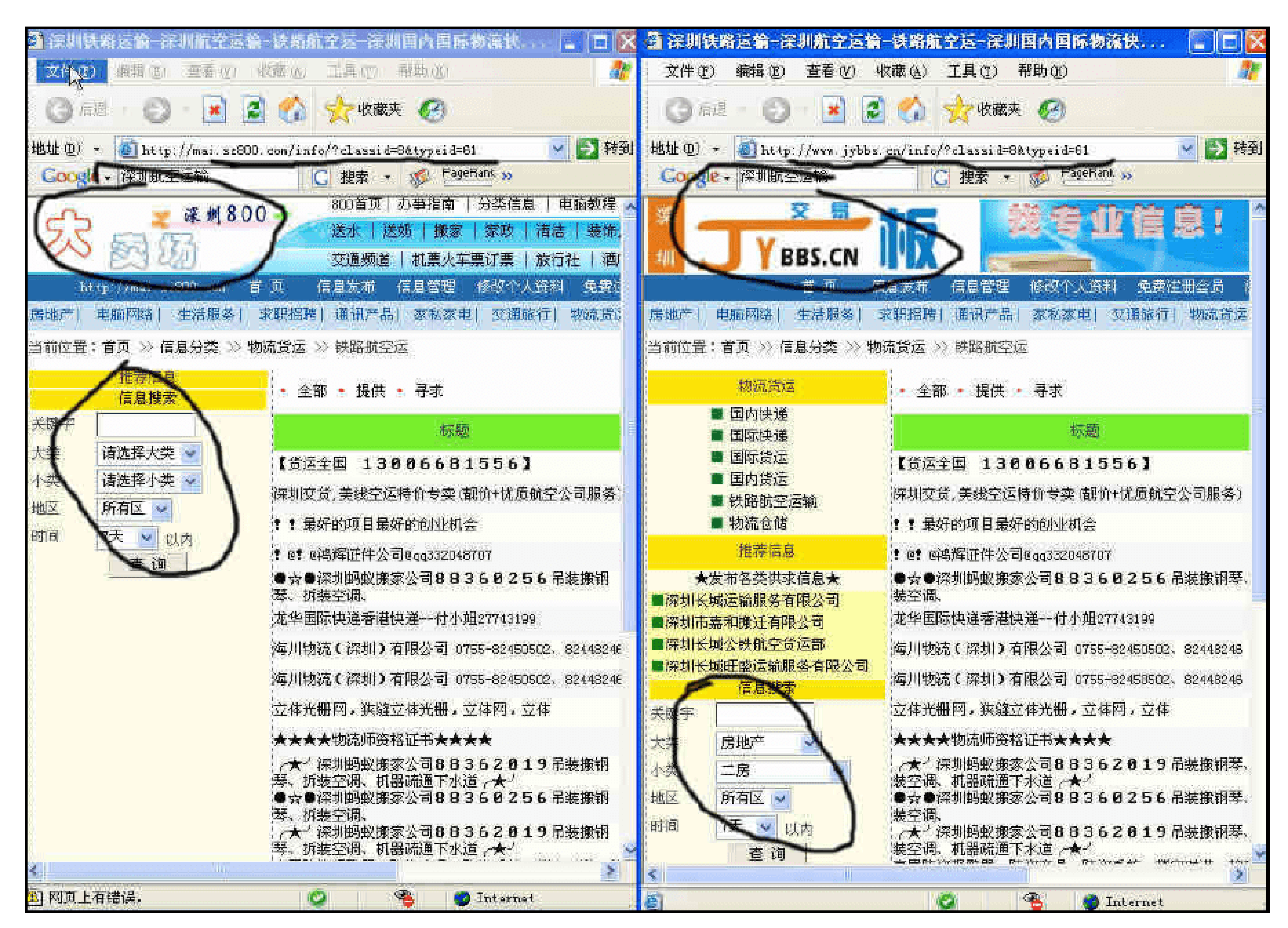


图 11-14 网站命名不同

深圳 160 对外大约有 120 多个这样互为镜向的页面新闻，要么就是内容完全一样只是路径不同，要么就是在域名、标题、字体、存放位置、结构上稍微调整。而网站的内容，似乎永远都没有变。第一个包含了第二个网站的内容，而第二个可能有包含了第三个网站的内容。反过来，第三个又有可能包含了前面两个网站的内容。120 多个网站相互关系错综复杂，集团做战，谁与争锋。

11.5.3 作弊带来的影响

深圳 160 的这种做法，实际上已经严重违反了搜索引擎的条款，也严重影响了访问者的感受，如果在 Google 里搜索深圳本地的信息，往往在前几页会看到好多他们网站的搜索结果，点过来点过去，好像搜索引擎已经完全被深圳 160 垄断了一样。

如以“深圳广告设计”为搜索，深圳 160 前 10 名有一个搜索结果外，第二页就有 5 个搜索结果是他们的，而且点进去的网站几乎都差不多，直接影响到了 Google 的正常使用。

有兴趣的读者可以随便搜索一些深圳的本地信息，如“深圳清洁”、“深圳装修设计”、“深圳包装”、“深圳运输”、“深圳印刷”等，相信你都可以感受到深圳 160 带给你的“超值享受”。

甚至可以这样说，在搜索深圳本地行业信息这个领域，除了少数几个热门行业，如“翻

译”、“鲜花”之外，深圳 160 几乎已经垄断了 Google 上关于深圳本地信息的搜索。Yahoo！虽然目前状况好一点，但仍然受到了其重大干扰。只有百度，百度毫不留情地删除了其绝大部分的域名。

Google 里搜索关键字“深圳广告设计”，搜索结果抓图，如图 11-15 所示。



图 11-15　谷歌搜索结果

11.5.4　搜索引擎将如何出招

面临这种大规模的，而且具有一定隐蔽性的，已经严重影响搜索引擎正常使用的作弊方法，搜索引擎该怎么办？Yahoo！和 Google 至今仍然没有发现或检测出来，只有百度，因为它们有专门的反垃圾小组存在，所以才能迅速地避免了被深圳 160 玩弄于股掌之间的尴尬。其他搜索引擎如 Yahoo！、中搜、搜狗等情况虽然也比较严重，但是比起 Google 来，已经算

是非常不错了。

由此可以看出，Google 在与有中国特色的各种作弊方式处理中，比起百度等本土搜索引擎来说，仍然有很大的距离。坚持不人工干涉的原则绝对没有错，关键是如何在坚持原则的基础上，怎么更好地改善搜索结果的质量，防止类似的被“劫持”的情况发生呢？Google 在中国反作弊要达到英文的水平，确实还有很长的路要走。

关于这种作弊方法，不知 Google 将会在什么时候采取行动。但是笔者永远相信，任何人，任何公司，谁影响了搜索引擎的正常使用，影响了访问者的感受，最终都会付出惨重的代价。

11.6 彩虹图书网

这个网站在百度、GG 及 Yahoo！上的收录页面都不错，Google上的收录大约为 36600 个；百度一下，找到相关网页约 79300 篇；Yahoo！上被收录的网页共有 17373 条。反向连接，Yahoo！查询出来的结果是链向该地址的网页共有 4281 条，但是发现很多连接都是自己的站，并且这些站是在一个服务器上，加上数据一样，这样的连接价值都不算高。

域名注册时间是：11-nov-2003，Google PR 值为 3.Sogou PR 值为 33。应该说这个域名还是比较有优势的，注册的时间比较早，这个站应该有潜力可挖。

1. 关于首页

顶部的<meta>标签：

```
<title>管理图书,企业管理图书,管理图书网,管理培训图书-彩虹图书网-首页</title>

<META cFONT-FAMILY:宋体; mso-ascii-font-family: Times New Roman'; mso-hansi-font-family: 'Times New Roman'">彩虹图书网是专业管理类图书网,主要经营管理培训教材,建筑图书,化工类书籍,电力图书,交通类图书,机械图书" name=description>

<META cFONT-FAMILY: 宋体; mso-ascii-font-family: 'Times New Roman'; mso-hansi-font-family: 'Times New Roman'">管理图书,企业管理图书,管理图书网,管理培训图书" name=keywords>
```

如果是企业站，做的是单一的产品，建议可以采用“关键词+企业名&网站名”来做 title 命名。但是书站不一样，关键词可以分不到具体的内容页和分类页面去，所有首页的写法可以参考如下：

```
<title>彩虹图书网-管理图书,企业管理,管理培训图书,专业的老牌管理图书网!</title>
```

description 标签写得不错，上面的“20 家出版社，300 多个城市”，这些数据很能说明问题。

关键词主要就是写上网站的主关键词，如“管理图书”百度第 1，GG 第 6。

网站 Logo 上面加上 alt 属性：<img src="/imagehe/top_001.gif" alt="管理图书" width="150" height="56">。

顶部中间的图片建议重新设计，加上促销、售后服务等具有诱惑性的因素，如在图片上加上“20 多个出版社合作，300 多个城市货到付款”等字样，给客户一个可信度。或者也可以加上一些广告标语：即日起，到彩虹图书网买书超过 200 元，可以获赠一本“资治通鉴”（当然不是送资治通鉴，可以送一些通用的书籍或者小礼品，这些也是营销中很重要的环节）。

右上角的地方，“书到付款”处可以考虑加上以下三个按钮。

- 收藏本站：便于客户把网站收藏起来，下次再看。
- 联系我们：把网站的联系方式加上去，电话、传真、手机、QQ、邮箱及论坛，把这些联系方式都放在一个页面。
- 购书帮助：这个连接到“付款方式、物流配送、注意事项”等页面。

这三个页面虽然很简单，但基本上是人们习惯的地方，加上右手鼠标的习惯，因此关注度是比较高的。

顶部的导航：可以增加 2 ~ 3 个互动性强的的栏目，如留言板、论坛等。增加和访客的互动和交流。

不要把 QQ 加在上面，这样显得很不专业，并不是说不放 QQ，可以加在右侧，做个跟随的客服，放在首页。加上常用的 QQ、MSN 等及时通讯软件，或者也可以用专业的客服软件，目前有几个免费的系统也不错。

首页底部：不定期的交换高质量并且相关度比较高的友情连接。加上网站备案信息、公司名称、联系方式等信息，提升访客的可信度，同时搜索引擎也很关注这些细节。

另外，底部的每个栏目点击进去必须要有实际的内容。而“网站首页”可以改成“关于我们”，写上彩虹图书的历史、公司简介等，能忽悠则忽悠，这个是门面的东西。

2. 关于内容页

列表页和栏目页问题不大，部分可以参考首页。

内容页主要就是关键词的布局和内容问题，搜索引擎是从上至下抓取的，对上面的和下面的内容比较关注。

顶部可以加上滚动的代码，如：

```
http://www.115book.com/Book29404
<marquee scrollAmount=2 width=780>
<td vAlign=center><nobr>你现在浏览的是<FONT color=red>《中国地产金牌户型大典》
</FONT>,<FONT color=red>《中国地产金牌户型大典》</FONT>面向全国多个城市提供货到付款
服务,<FONT color=red>《中国地产金牌户型大典》</FONT>是专业书籍,收到<FONT
color=red>《中国地产金牌户型大典》</FONT>当面付款。让你真正放心购买<FONT color=red>
```

```
中《中国地产金牌户型大典》</FONT>。</nobr></td></marquee>。
```

“首页”>“图书分类”>“书名”，这个地方都要加上链接，并且图书分类最好具体到详细的分类，便于用户返回。

书的介绍区用黄色大字号来显示书名，如《中国地产金牌户型大典》，而不是“书名：中国地产金牌户型大典”，直接把这个关键词呈现给搜索引擎。

可以考虑加上书名号，因为很多客户搜索书名的时候也会带上“《》”，而且加上书名号对搜索不带书名号的书名并没有影响。不带书名的结果会多很多，他们的排名还是有些许区别的。

书中的图片可以加上 alt 属性，描述直接就是书名，通过多种途径来说明这个关键词是很重要的内容。

内容中可以出现 1 ~ 2 次关键词，但是要很自然的融合进去，不能直接的堆积起来，这个很容易被搜索引擎认为是恶意的关键词堆积作弊。

底部的“未经许可不得复制：中国地产金牌户型大典”以及“欢迎你对中国地产金牌户型大典发表评论，你所发表的对中国地产金牌户型大典的一切言论，仅代表你个人观点与本站无关。你还可以购买本站中国地产金牌户型大典类似以下图书：”。这个可以保留，但是最好是加上颜色，或者加粗以示强调。

3. 书目介绍的部分

书名可以使用链接的形式，锚文本就是书名，这样也可以起到强调关键词的作用。<title>中国地产金牌户型大典-2007 年 10 月-彩虹图书网</title>，不知道这个通过出版日期来的流量多不多，如果多的话可以保留，同时建议加上栏目的分类。

4. 在线帮助

http://www.115book.com/help.asp，这个环节很重要，这里面的购物流程、付款方式、运输方式、商品销售和售后服务、常见问题、关于我们、注册条约、法律和版权声明、交易条款最好不要完全一样，把每一个都写清楚，让顾客一目了然。

<meta>标签的书写如下：

```
<title>在线帮助-彩虹图书网-专业的管理培训,管理培训教材,管理培训图书网站</title><meta
name="Description" cFONT-FAMILY: 宋体; mso-ascii-font-family: 'Times New
Roman'; mso-hansi-font-family: 'Times New Roman'">专业的管理培训,管理培训教材,
管理培训图书网站,欢迎订购,电话：010-87654321" />。

<meta name="Keywords" cFONT-FAMILY: 宋体; mso-ascii-font-family: 'Times New
Roman'; mso-hansi-font-family: 'Times New Roman'">管理培训,管理培训教材,管理培
训图书,智能培训" />。
```

5. 意见反馈

http://www.115book.com/from/index.asp，这个地方顶部的图片全部都不显示，这对用户

的体验存在很大的影响。所以最好加上留言本，让客户可以留言，管理员可以回复，便于互动。如果客户想订购站上没有的书，也可以及时反馈上来。

除此之外，可以加上论坛进行互动，论坛可以让客户提供信息，一个随时更新的网站，搜索引擎是很喜欢的，谁都不喜欢一个站长长期不更新。论坛可以聚集人气，可以有效地留住访客，同时给大家建立一个交流的平台。

6. 外部连接和推广策略

与同行业网站或者相关行业的网站交换友情连接。如培训站，管理资讯站等。在行业门户性的网站上发布信息，带上网站链接，因为这些网站权重都比较高，可以利用他们的优势。在比较大的，人气比较高的网站上发布信息，但是不能直接 AD，可以以介绍新书的形式发布，然后带上我们自己站的链接。

不宜直接出现网址的地方，可以使用链接，如：彩虹图书网。或者直接以书名即关键词为锚文本，如：《中国地产金牌户型大典》。

我们还可以在大的论坛签名上加上这些链接，一般的论坛都会有这样的功能。可以加上网站的链接，想推新书的时候可以换成新书的链接。

利用 Blog，现在各大门户和行业性的网站都有 Blog，可以申请，然后在上面发布一些书的文章，并带上具体页面的连接地址。同时在友情链接上也可以挂上主站的链接，经常性地进行更新维护，一般 Blog 的权重都比较高。

利用软文推广，软文就是提供给大家一定信息，但是其目的就是用这些信息来推广自己的站，这样让大家来转载，转载的人多了，文章中提到的外链也会跟着多，外链对网站排名的影响还是很重要的。同时，软文的发布地址可以是小说站、书城论坛、读书频道等地方，因为这个部位的相关度很高，并且很多都是潜在访客。

11.7 小结

搜索引擎优化通俗地说其实就是把握好很多网站设计制作中的因素，把站内的内容通过对搜索引擎友好的手段呈现给搜索引擎，同时也呈现给访客。这中间有一个共同点，就是搜索引擎面对的用户也是搜索者，利于访客的用户体验的站也同样更容易得到搜索引擎的青睐。但是这些都是一个长期的、循序渐进的过程，中间需要不断地改进才能达到在搜索引擎中排名上升效果，从而带来有效流量，赢得订单。